

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 106.239 Seconds
(without alignments)
1870.310 Million cell updates/sec

Title: US-09-785-215-2
Perfect score: 4058
Sequence: 1 MLPGLALLLLAAWTARALEV.....KMQQNGYENPTYKFFEQMQN 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match	Length	DB	ID	Description

1	3981	98.1	770	6	Q9TUI0	Q9tui0 sus scrofa
2	3719.5	91.7	751	13	Q9DGJ7	Q9dgj7 gallus gall
3	3522.5	86.8	695	11	Q60496	Q60496 cavia sp. p
4	3489.5	86.0	695	11	P97487	P97487 mus musculu
5	3403.5	83.9	747	13	Q91963	Q91963 xenopus. ap
6	3377.5	83.2	695	13	Q9DGJ8	Q9dgj8 gallus gall
7	3163.5	78.0	693	13	Q98SG0	Q98sg0 xenopus lae
8	3152.5	77.7	695	13	Q98SF9	Q98sf9 xenopus lae
9	3065	75.5	607	11	Q99K32	Q99k32 mus musculu
10	2914	71.8	699	13	O57394	O57394 narke japon
11	2785.5	68.6	738	13	Q90W28	Q90w28 brachydanio
12	2763.5	68.1	780	13	O73683	O73683 tetraodon f
13	2717	67.0	569	13	Q9PVL1	Q9pvl1 gallus gall
14	2699	66.5	737	13	O93279	O93279 fugu rubrip
15	2562.5	63.1	534	13	O93296	O93296 gallus gall
16	2437	60.1	694	13	Q8UUR9	Q8uur9 brachydanio
17	2288.5	56.4	612	13	Q9I9E7	Q9i9e7 brachydanio
18	1960.5	48.3	763	11	Q61482	Q61482 mus musculu
19	1956.5	48.2	751	11	Q60709	Q60709 mus musculu
20	1928	47.5	384	11	Q8BPC7	Q8bpc7 mus musculu
21	1733.5	42.7	695	4	Q13861	Q13861 homo sapien
22	1721	42.4	669	4	Q14662	Q14662 homo sapien
23	1709.5	42.1	695	11	Q64348	Q64348 mus musculu
24	1604.5	39.5	472	13	Q8UUS0	Q8uus0 brachydanio
25	1350.5	33.3	357	13	Q8UUI8	Q8uui8 brachydanio
26	1302	32.1	523	4	Q14594	Q14594 homo sapien
27	1267	31.2	522	4	Q9BT36	Q9bt36 homo sapien
28	1090	26.9	218	11	Q8BPV5	Q8bpv5 mus musculu
29	788	19.4	160	11	Q9QZ78	Q9qz78 cavia sp. p
30	771	19.0	239	13	Q8UUI7	Q8uui7 brachydanio
31	678	16.7	136	6	P79307	P79307 sus scrofa
32	577	14.2	113	13	Q8JH58	Q8jh58 chelydra se
33	561	13.8	182	11	Q9CYS4	Q9cys4 mus musculu
34	478	11.8	97	6	Q28673	Q28673 oryctolagus
35	393.5	9.7	82	4	Q16019	Q16019 homo sapien
36	389.5	9.6	82	4	Q16014	Q16014 homo sapien
37	387.5	9.5	82	4	Q16020	Q16020 homo sapien
38	376	9.3	79	11	O35463	O35463 cricetus
39	374	9.2	74	11	Q60495	Q60495 cavia sp. a
40	335	8.3	208	11	Q8R0R7	Q8r0r7 mus musculu
41	239	5.9	49	6	O97917	O97917 bos taurus
42	207	5.1	3198	5	Q9U8G8	Q9u8g8 manduca sex
43	205.5	5.1	759	5	Q8IT91	Q8it91 ancylostoma
44	194.5	4.8	2225	5	O45881	O45881 caenorhabdi
45	187	4.6	3060	5	Q9VAV4	Q9vav4 drosophila

ALIGNMENTS

RESULT 1

Q9TUI0

ID Q9TUI0 PRELIMINARY; PRT; 770 AA.

AC Q9TUI0;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Amyloid precursor protein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimura A., Takahashi T.;
RT "Amyloid Precursor Protein 770.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB032550; BAA84580.1; -.
DR HSSP; P05067; 1AAP.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00006; A4_EXTRA; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.
KW Protease inhibitor; Serine protease inhibitor.
SQ SEQUENCE 770 AA; 86961 MW; 5F7A1DCB2BCC583E CRC64;

Query Match 98.1%; Score 3981; DB 6; Length 770;
Best Local Similarity 97.8%; Pred. No. 2e-241;
Matches 753; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
		: : :	
Db	1	MLPGLALVLLAAWTARALEVPTDGNAGLLAEPQVAMFCGKLNMHMNVQNGKWESDPSGTK	60
Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG	120
Db	61	TCIGTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRSRKQCKTHTHIVIPYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
		: :	
Db	181	GVEFVCCPLAEESDNIDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVADVEEE	240
Qy	241	EADDDDEDEDGDEVEEEAEPEYEEATERTTSIATTTTTTTTESVEEVVREVCSEAETGPC	300
		:	
Db	241	EAEDDEDEDGDEVEEEAEPEYEEATERTTSIATTTTTTTTESVEEVVREVCSEAETGPC	300
Qy	301	RAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
		:	
Db	301	RAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSVMSQSLLKTTQEHL PQD	360

Qy 361 PVKLPPTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
 |||||
 Db 361 PVKLPPTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
 |||||
 Qy 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480
 |||||
 Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480
 |||||
 Qy 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 |||||
 Db 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 |||||
 Qy 541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET 600
 |||||
 Db 541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET 600
 |||||
 Qy 601 KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 |||||
 Db 601 KTTVELLPVNGEFSLDDLQPWHPFGVDSVPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 |||||
 Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 |||||
 Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 |||||
 Qy 721 VMLKKKQYTSIHGHVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 |||||
 Db 721 VMLKKKQYTSIHGHVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 |||||

RESULT 2

Q9DGJ7

ID Q9DGJ7 PRELIMINARY; PRT; 751 AA.
 AC Q9DGJ7;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Beta-amyloid precursor protein 751 isoform.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sarasa M., Rodolosse A., Sorribas V.;
 RT "Cloning of full-length chicken beta-amyloid precursor protein
 RT isoforms.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF289219; AAG00594.1; -.
 DR HSSP; P05067; 1BA4.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR Pfam; PF00014; Kunitz_BPTI; 1.


```

Db      580 ETKTTVELLPVDGEFSLDDLQPWHPFGVDSVPANTENEVEFPVDARPAADRGLTTRPGSGL 639
Qy      659 TNIKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVVIAIVIVI 718
Db      640 TNVKTEEVSEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVVIAIVIVI 699
Qy      719 TLVMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEOMQN 770
Db      700 TLVMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEOMQN 751

```

RESULT 3

Q60496

```

ID   Q60496          PRELIMINARY;          PRT;          695 AA.
AC   Q60496;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative amyloid precursor protein.
OS   Cavia sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10143;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=97236426; PubMed=9116031;
RA   Beck M., Mueller D., Bigl V.;
RT   "Amyloid precursor protein in Guinea pigs - complete cDNA sequence and
RT   alternative splicing.";
RL   Biochim. Biophys. Acta 1351:17-21(1997).
DR   EMBL; X97631; CAA66230.1; -.
DR   HSSP; P05067; 1BA4.
DR   InterPro; IPR001868; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
SQ   SEQUENCE 695 AA; 78701 MW; 5196A0C4017F16AB CRC64;

```

```

Query Match          86.8%;  Score 3522.5;  DB 11;  Length 695;
Best Local Similarity 88.2%;  Pred. No. 1e-212;
Matches 679;  Conservative 7;  Mismatches 9;  Indels 75;  Gaps 1;

```

```

Qy      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSTGK 60
Db      1  MLPSLALLLLTTWTARALEVPTDGNAGLLAEPQIAMFCGKLNHMHMNVQNGKWEPTDPSGK 60
Qy      61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
Db      61  TCIGSKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRSRKQCKTHPHFVIPYRCLVG 120
Qy      121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

```

Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Db	181	GVEFVCCPLAEESDNIDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVADVEEE	240
Qy	241	EADDDDEDEDGDEVEEEAEPEYEEATERTTSIATTTTTTTTSESVEEVVREVCSEQAETGPC	300
Db	241	EADDDDEDVEDGDEVEEEAEPEYEEATEKTTTSIATTTTTTTTSESVEEVVR-----	288
Qy	301	RAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	289	-----	288
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Db	289	---VPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	345
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	346	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	405
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	406	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	465
Qy	541	MNQSLSLLYNVPVAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET	600
Db	466	MNQSLSLLYNVPVAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET	525
Qy	601	KTTVELLPVNGEFSLLDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	526	KTTVELLPVNGEFSLLDLQPWHPFGVDSVPANTENEVEPVDARPAADRGLTTRPGSGLTN	585
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	586	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	645
Qy	721	VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQRN	770
Db	646	VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQRN	695

RESULT 4

P97487

ID P97487 PRELIMINARY; PRT; 695 AA.
AC P97487; P97942;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hippocampal amyloid protein.
GN APP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	346	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	405
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	406	QAVPPRPHHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	465
Qy	541	MNQSLSLLYNVPVAVAEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNDALMPSLTET	600
Db	466	MNQSLSLLYNVPVAVAEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNDALMPSLTET	525
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDPARPAADRGLTTRPGSGLTN	660
Db	526	KTTVELLPVNGEFSLDDLQPWHPFGVDSVPANTENEVEPVDPARPAADRGLTTRPGSGLTN	585
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAITVIVITL	720
Db	586	IKTEEISEVKMDAEFGHDSGFVVRHQKLVFFAEDVGSNKGAIIGLMVGGVVIAITVIVITL	645
Qy	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770
Db	646	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	695

RESULT 5

Q91963

ID Q91963 PRELIMINARY; PRT; 747 AA.
AC Q91963;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE APP747.
GN APP747.
OS Xenopus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae.
OX NCBI_TaxID=8353;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93129227; PubMed=1282805;
RA Okado H., Okamoto H.;
RT "A Xenopus homologue of the human beta-amyloid precursor protein:
RT developmental regulation of its gene expression.";
RL Biochem. Biophys. Res. Commun. 189:1561-1568(1992).
DR EMBL; S52417; AAB24853.1; -.
DR HSSP; P05067; 1HZ3.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 1.


```

Db      462 ERMNQSFSLLYKVPAAVEEIQDEVDELFOKEQNYSDDMVSNMVS DHRVSYGNDALMP SLS 521
Qy      599 ETKTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGL 658
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      522 ETKTTVELLPVDGEFNI EDLQPWHSFGVDSVPANTENEVEPVDARPAADRGLTTRPGSGL 581
Qy      659 TNIKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVI 718
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      582 TNIKTEEISEVKMDSEYRHDTAYEVHHQKLVFFAEEVGSNKGAIIGLMVGGVVIATVIVI 641
Qy      719 TLVMLKKKQYTSIHG VVEVDAAVTPEERHLSKMQQNGYENPTYKFFE QMQN 770
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      642 TLVMLKKKQYTTIHG VVEVDAAVTPEERHLTKMQQNGYENPTYKFFE QMQN 693

```

RESULT 8

Q98SF9

```

ID      Q98SF9          PRELIMINARY;          PRT;          695 AA.
AC      Q98SF9;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      Beta-amyloid precursor protein B.
GN      APP.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC      Xenopodinae; Xenopus.
OX      NCBI_TaxID=8355;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Van den Hurk W.H.;
RL      Thesis (2001), Department of Biological Sciences,
RL      University of Nijmegen, Nijmegen, Netherlands.
DR      EMBL; AJ298151; CAC37194.1; -.
DR      HSSP; P05067; 1HZ3.
DR      InterPro; IPR001868; A4_APP.
DR      InterPro; IPR001255; Beta-APP.
DR      Pfam; PF02177; A4_EXTRA; 1.
DR      Pfam; PF03494; Beta-APP; 1.
DR      PRINTS; PR00203; AMYLOIDA4.
DR      SMART; SM00006; A4_EXTRA; 1.
DR      PROSITE; PS00319; A4_EXTRA; 1.
DR      PROSITE; PS00320; A4_INTRA; 1.
KW      Signal.
FT      SIGNAL          1          18          POTENTIAL.
SQ      SEQUENCE      695 AA; 78803 MW; DC14EB02AFB0204A CRC64;

```

```

Query Match          77.7%; Score 3152.5; DB 13; Length 695;
Best Local Similarity 78.9%; Pred. No. 1.6e-189;
Matches 609; Conservative 39; Mismatches 45; Indels 79; Gaps 5;

```

```

Qy      1 MLPGIALLLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMH MNVQNGKWDSDPSGTK 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MLPHITLLVLTA-GALALEVPADGNGGLLAEPQIAMFCGKLNMH MNVQNGKWETDVS GTK 59
Qy      61 TCIDTKEGILQYCQEVYP ELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

```


Qy 524 AQIRSQVMTHLRVIYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISE 583
 Db 361 AQIRSQVMTHLRVIYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISE 420

Qy 584 PRISYGNDALMPSLTETKTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPV DAR 643
 Db 421 PRISYGNDALMPSLTETKTTVELLPVNGEFSLDDLQPWHPFGVDSVPANTENEVEPV DAR 480

Qy 644 PAADRGLTTRPGSGLTNIKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAI I 703
 Db 481 PAADRGLTTRPGSGLTNIKTEEISEVKMDAEFGHDSGFVHRHQKLVFFAEDVGSNKGAI I 540

Qy 704 GLMVGGVVIATVIVITLVMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYK 763
 Db 541 GLMVGGVVIATVIVITLVMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYK 600

Qy 764 FFEQMQN 770
 Db 601 FFEQMQN 607

RESULT 10

O57394

ID O57394 PRELIMINARY; PRT; 699 AA.
 AC O57394;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE EL amyloid precursor protein 699.
 GN EL APP699.
 OS Narke japonica (Electric ray).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
 OC Torpediniformes; Narcinoidei; Narkidae; Narke.
 OX NCBI_TaxID=62965;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Electric lobe;
 RX MEDLINE=98129705; PubMed=9461486;
 RA Iijima K., Lee D.-S., Okutsu J., Tomita S., Hirashima N., Kirino Y.,
 RA Suzuki T.;
 RT "cDNA isolation of Alzheimer's amyloid precursor protein from
 RT cholinergic nerve terminals of the electric organ of the electric
 RT ray.";
 RL Biochem. J. 330:29-33(1998).
 DR EMBL; AB005544; BAA24230.1; -.
 DR HSSP; P05067; 1HZ3.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 SQ SEQUENCE 699 AA; 78879 MW; 952915C309D50E5C CRC64;

Query Match 71.8%; Score 2914; DB 13; Length 699;
Best Local Similarity 72.9%; Pred. No. 1.5e-174;
Matches 568; Conservative 59; Mismatches 58; Indels 94; Gaps 9;

```
Qy      2 LPG-LALILLAAWTA-----RALEVPTDGNAGLL-AEPQIAMFCGRILNMHMNVQNGKW 52
      ||| | : |||| | ||||| ||| ||||| ||| : ||| : ||| |||
Db      5 LPGRILGMLLLAAAAALVLAPLCRALEVPTDGGAGLLAAEPQIAMFCGKILNMHVNVQTGKW 64

Qy     53 DSDPSGKTCTCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFV 112
      ||||| || ||||| : ||||| : ||||| ||||| : ||||| ||| |
Db     65 VSDPSGTNTCFGTKEGILRYCQEVYPDLQITNVVEANQPITIQNWCKKGRKQCKGHPHIV 124

Qy    113 IPYRCLVGEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLL 172
      : ||||| ||||| ||||| ||||| : ||| ||| : ||||| ||| : ||||| |||
Db    125 VPYRCLVGEFVSDALLVPDKCKFLHREKMDTCESHLYWHTVAKETCGDKIMNLHDYGMLL 184

Qy    173 PCGIDKFRGVFEVCCPLAEESDNVDSADAEEDSDVWWGGADTDYADGSEDKVVEVAEEE 232
      ||||| : ||||| ||||| : ||| : ||| : ||||| ||| ||||| ||| |||
Db    185 PCGIDEFRGVFEVCCPIPEENDKIDS-DMDEEDSDVWWGGDDADYADGG-DKTV----EE 238

Qy    233 EVAEVEEEEEADDDDEDEDGDEVEEEE-AEPEYEEATERTTSIATTTTTTTESVEEVVREVC 291
      : | |||| : | |||| |:::| | : ||| ||| : ||||| : |||||
Db    239 KPIEEEEEEDESIDDEDDDDLDDEVVEDQYEDPTEHTTS---STTTTTEAIEEVVR--- 292

Qy    292 SEQAETGPCRAMISRWFYDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLK 351
Db    293 ----- 292

Qy    352 TTQEPLARDPVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMR 411
      : ||||| ||||| ||||| ||||| : ||||| ||||| ||||| : |||
Db    293 -----VPTTAASTPDAVDKYLETPGDENEHAYFQKAKERLEAKHRERMSKIMR 340

Qy    412 EWEEAERQAKNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRL 471
      ||||| ||||| ||||| ||||| ||||| ||||| : ||||| ||||| ||||| :
Db    341 EWEEAERQAKNLPKADKKAVIQRFQEMVESLEQEAAASERQQLVETHMARVEAMLNDRRI 400

Qy    472 ALENYITALQAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVM 531
      ||||| : ||||| ||||| ||||| ||||| ||||| : ||||| ||||| : |||||
Db    401 ALENYLAALQADPPRPRHVLNALKKYSRAEQKDRQHTLKHFDHVRVADPEKAAQIKSQVM 460

Qy    532 THLRVIYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGN 591
      ||| ||| ||||| ||||| ||||| ||||| : ||||| ||||| : ||||| |||||
Db    461 THLHVIDERMNQSLSLLYKVPVAEEIQDEVDELLQRESYMDDMMANSVSDTRISYGN 520

Qy    592 ALMPSLTETKTTVELLPVNGEFLDDLQPWHSFGADSVANTENEVEPVDPARPAADRGLT 651
      ||| : ||| : ||||| : ||||| : ||| ||||| ||| : ||| ||||| ||||| |||||
Db    521 ALVPSLSETKTTIELLPDDGEFILDDLQPPHPFVIESIPANTENEVEPVDPARPAPDRGLT 580

Qy    652 TRPGSGLTNIKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV 711
      ||||| ||||| : ||| : ||| : ||||| ||||| ||||| : ||||| ||||| |||||
Db    581 TRPGSGLTGIKTEEIAELKMETEFQQDSGYEVHHQKLVFFPKDVGSNKGAIIGLMVGGVV 640

Qy    712 IATVIVITLVMLKKKQYTSIHHGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQM 770
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    641 IATVIVITLVMLKKKQYTSIHHGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQM 699
```

RESULT 11

Q90W28

ID Q90W28 PRELIMINARY; PRT; 738 AA.
AC Q90W28;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Amyloid precursor protein.
GN APPA OR APP.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Groth C., Lardelli M.;
RT "Expression analysis of zebrafish app.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF389401; AAK64495.1; -.
DR ZFIN; ZDB-GENE-000616-13; appa.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00006; A4_EXTRA; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW Protease inhibitor; Serine protease inhibitor.
SQ SEQUENCE 738 AA; 83577 MW; AF480F6D308FD298 CRC64;

Query Match 68.6%; Score 2785.5; DB 13; Length 738;
Best Local Similarity 69.9%; Pred. No. 1.9e-166;
Matches 540; Conservative 85; Mismatches 101; Indels 47; Gaps 14;

Qy 5 LALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGKTCTCID 64
| :||:| : ||||:| ||||| ||||| :|||:| :|||: ||||:| :||
Db 6 LFILLMAVASTLAVEVPSDSGTGLLAEPQIAMFCGKLNMHINIQSGKWEPPDSGSKSCIG 65
Qy 65 TKEGILQYCQEVYPQLQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVGEFVS 124
||||| ||||| ||||| :| :|||: ||||:| | :||| |||||
Db 66 NKEGILQYCQEVYPQLQITNVVEANQPVSIWDWCKKSRKQCRSHMHIVVPYRCLVGEFVS 125
Qy 125 DALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFRGVEF 184
||||| ||||| ||||| :||: ||||| ||||:| :|| ||||| ||||| :|||
Db 126 DALLVPDKCKFLHQERMDMCESHLHWHTVAKESCGDRSMNLHDYGMLLPCGIDRFRGVEF 185
Qy 185 VCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEEEEEEADD 244
|||| | : || ||||| ||||| : || : | : : | | | :||:| :||

Db	186	VCCP-ADAGKESESAAVEEDSDVWVGGAADY TENSMT R--DAAAEP AVLE-DEDEADE	241
Qy	245	DED-DEGD-----EVEEEAE E P YEEATERTT-SIATTTTTTTTESV E EVVREVCSEQAET	297
Db	242	EEDEDQDGDGRDEKIEEEEE--EERTQSTSAALTSTTTTTTTTESV E EVVREVCFA SAET	299
Qy	298	GPCRAMISRWFYFDVTEGKCAPFFYGGCGGNRN NFDTEEYCMAVCGSAMSQSLLKTTQEPL	357
Db	300	GPCRAMLSRWYYVREERRCAPFIYGGCGGNRN NFESE EYCLSVCSGV-----	346
Qy	358	ARDPVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAE	417
Db	347	-----LPTPSSSPDAVDRYLETPADENEHAHFLKAKESLETKHRERMSQVMREWEEAE	400
Qy	418	RQAKNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYI	477
Db	401	RQAKSLPRNDKKAVIQHFQEKVEALEQESASERQQLVETHMARVEALLNDRRRRLALESYL	460
Qy	478	TALQAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVI	537
Db	461	SALQADPPRPRHVSLLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRPQVLTHLRVI	520
Qy	538	YERMNQSLSLLYNVPAVAEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSL	597
Db	521	EERMNQSLGLLYKVPGVADDIQDQV-ELLQREQQEMSAQLANLQSDARVSYGNDALMPDS	579
Qy	598	TETKTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPV DARPAADRGLTTRPGSG	657
Db	580	T---AGLELLPAEDTQGF GFIHP-ESFN---QP NTHNQVEPV DARVPDLDLATRPVSG	631
Qy	658	LTNIKTEEISEVKMDAEFRHDSGYEVHHQKL VFFAEDVGSNKGAIIGLMVGGVVIATVIV	717
Db	632	L---KPDDIPELRMEAEERHS---EVYHQKL VFFAEDVSSNKGAIIGLMVGGVVIATIIV	685
Qy	718	ITLVMLKKKQYTSIH HGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEOMQN	770
Db	686	ITLVMLRK KQYTSIH HGIIEVDAAVTPEERHLSKMQQNGYENPTYKFFEOMHN	738

RESULT 12

ID O73683 PRELIMINARY: PRT: 780 AA.

DT 01-AUG-1998 (TrEMBLrel. 07, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Alzheimer's disease amyloid A4 protein homolog precursor [Contains:

DE beta-amyloid protein (Beta-APP) (A-beta)].

GN APP.

OS Tetraodon fluviatilis (Puffer fish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

OC Tetradontoidea; Tetraodontidae; Tetraodon.

OX NCBI_TaxID=47145;

RN [1]

RP SEQUENCE FROM N.A.

Qy 67 EGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVGEFVSDA 126
 Db 68 EGILQYCQEVYPELQITNVVEANQPVS IQNWCKKGRKQCRSHMHIVVPYRCLVGEFVSDA 127

Qy 127 LLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFRGVEFVC 186
 Db 128 LLVPDKCKFLHQERMNQCESHLHWHTVAKESCGDRAMNLHDYGMLLPCGIDRFRGVEFVC 187

Qy 187 CPLAEESDNVDSADAEEDSDVWWGGADTDYADGS-----EDKVVEVAEEEE 232
 Db 188 CP-AEAERDMDSTEKDADDSDVWWGGADNDYSDNSMVREPEPAEQEETRPSVVEEEEEEG 246

Qy 233 EVAEVEEEEE-----ADDEDEDGDEVEEEAEPEYEEATERTTTSIA 273
 Db 247 EVAQEDDEEEEEEVLDTDQGDGEEDHEAADDEEEEDVDEIDAFGESDDVDADEPTTNVA 306

Qy 274 ---TTTTTTTESVEEVVREVCSEQAETGPCRAMISRWFVDVTEGKCA-PFFYGGCGGNRN 329
 Db 307 MTTTTTTTTTESVEEVVRMFCWAHADTGPCTASMPSWYFDAVDGRMTMYELMYGGCGGNMN 366

Qy 330 NFDTEEYCMVCGSAMSQSLLKTTQEPLARDPVKLPPTAASTPDAVDKYLETPGDENEHA 389
 Db 367 NFEESEYCLSVCSV-----VPTDMPSSPDAVDHYLETPADENEHA 407

Qy 390 HFQKAKERLEAKHRERMSQVMREWEAAERQAKNLPKADKKAVIQHFQEKVESLEQEAANE 449
 Db 408 HFQKAKESLEAKHRERMSQVMREWEAAERQAKNLPKADKKKIVIQRFQEKVEALEQEAASE 467

Qy 450 RQQLVETHMARVEAMLNDRRRLALENYITALQAVPPRPRHVFNMLKKYVRAEQKDRQHTL 509
 Db 468 RQQLVETHMARVEALLNDRRRLALENYLTALQQDPPRPRHVFSLKKYVRAEQKDRQHTL 527

Qy 510 KHFEHVRMVDPKKAAQIRSQVMTHLRVIYERMNQSLSLLYNVPAAVEEIQDEVDELLQKE 569
 Db 528 KHFEHVRMVDPKKAAQIRPQVLTHLRVIEERMNQSLGLLYKVPGVADDIQDQV-ELLQRE 586

Qy 570 QNYSDDVLANMISEPRISYGNDALMPSLTETKTTVELLPVNGEFSLDDLQPDWH--SFGAD 627
 Db 587 QAEMAQQLANLQTDVRVSYGNDALMPDQELGDGQADLLP--QEDTLGGVGVFVHPESFN-- 642

Qy 628 SVPANTENEVEPVDPARPAADRGLTTRPGSGLTNIKTEEISEVKMDAEFRHDSGYEVHHQK 687
 Db 643 --QLNTENQVEPVDSRPTFERGVPTRP---VTGKSMEAVPELRMETEDRQSTEYEVHHQK 697

Qy 688 LVFFAEDVGSNKGAIIGLMVGGVVIATVIVITLVMLKKKQYTSIHHGVVEVDAAVTPEER 747
 Db 698 LVFFAEDVGSNKGAIIGLMVGGVVIATVIVITLVMLRKKQYTSIHHGII EVDAAVTPEER 757

Qy 748 HLSKMQQNGYENPTYKFFEQMQN 770
 Db 758 HLSKMQQNGYENPTYKFFEQMQN 780

RESULT 13

Q9PVL1

ID Q9PVL1 PRELIMINARY; PRT; 569 AA.
 AC Q9PVL1;

QY 484 PPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYERMNQ 543
 |||
 Db 284 PPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAVQIRSQVMTHLRVIYERMNQ 343
 |||
 QY 544 SLSLLYNVPAVAEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTETKTT 603
 |||
 Db 344 SLSFLYNVPAVAEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPTLTETKTT 403
 |||
 QY 604 VELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTNIKT 663
 |||
 Db 404 VELLPVVGEFSLDDLQPWHPFGVDSVPANTENEVEPVDARPAADRGLTTRPGSGLTNVKT 463
 |||
 QY 664 EEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITLVML 723
 |||
 Db 464 EEVSEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIANVIVITLVML 523
 |||
 QY 724 KKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 |||
 Db 524 KKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 569
 |||

RESULT 14

093279

ID 093279 PRELIMINARY; PRT; 737 AA.
 AC 093279;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein homolog precursor [Contains:
 DE beta-amyloid protein (Beta-APP) (A-beta)].
 GN APP.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Takifugu.
 OX NCBI_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98252138; PubMed=9599080;
 RA Villard L., Tassone F., Crnogorac-Jurcevic T., Clancy K., Gardiner K.;
 RT "Analysis of pufferfish homologues of the AT-rich human APP gene."
 RL Gene 210:17-24(1998).
 CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
 CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
 CC G(O) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION
 CC WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC
 CC RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE
 CC NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF
 CC PHOSPHORYLATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 PROTEASE INHIBITOR DOMAIN BELONGING TO THE
 CC BPTI/KUNITZ FAMILY OF INHIBITORS.
 DR EMBL; AF090120; AAD13392.1; -.
 DR HSSP; P05067; 1H23.

DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; FALSE_NEG.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.
 KW Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;
 KW Serine protease inhibitor.

FT	SIGNAL	1	18	POTENTIAL.
FT	CHAIN	19	737	ALZHEIMER'S DISEASE AMYLOID A4 PROTEIN
FT				HOMOLOG.
FT	CHAIN	639	681	BETA-AMYLOID PROTEIN (POTENTIAL).
FT	DOMAIN	19	668	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	669	689	POTENTIAL.
FT	DOMAIN	690	737	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	286	344	BPTI/KUNITZ INHIBITOR.
FT	SITE	726	729	CLATHRIN-BINDING (BY SIMILARITY).
FT	ACT_SITE	300	301	REACTIVE BOND.
FT	DISULFID	290	340	BY SIMILARITY.
FT	DISULFID	299	323	BY SIMILARITY.
FT	DISULFID	315	336	BY SIMILARITY.
FT	CARBOHYD	522	522	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	737 AA;	82856 MW;	6FAD01E2E3B2B7E2 CRC64;

Query Match 66.5%; Score 2699; DB 13; Length 737;
 Best Local Similarity 67.2%; Pred. No. 5e-161;
 Matches 519; Conservative 93; Mismatches 110; Indels 50; Gaps 12;

Qy	7	LLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGKTCTCIDTK	66
		: : : : : : : : ::	
Db	8	VLLLVATLTRSSEIPADDTVGLLTPEQVAMFCGKLNMHINVQNGKWESDPSGKTSCLNTK	67
Qy	67	EGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVGEFVSDA	126
Db	68	EGILQYCQEVYPELQITNVVEANQPVSIQNWCKKGRKQCRSHTHIVVPYRCLVGEFVSDA	127
Qy	127	LLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFRGVEFVC	186
Db	128	LLVPDKCKFLHQERMNQCESHLHWHTVAKESCGDRSMNLHDYGMLLPCGIDRFRGVKFC	187
Qy	187	CPLAEESDNVDSADAEEDSDVWWGGADTDYADGS---EDKVVEVAEEEEVAEVEEEEEAD	243
		:: : : : : : :	
Db	188	CP-AETEQETDSSEVEGEESDVWWGGADPEYSSENSPPTPSRATYVAGD---AFERDENG	243
Qy	244	DDEDDDEDGDEVEEEAEPEYEEATERTTSIA--TTTTTTTSESVEEVVREVCSEQAETGPCR	301
		: : : : : :	
Db	244	GDEDEDDDEDVDPTDE---QESDERTANVAMTTTTTTTTTSESVEEVVRAVCWAQAESGPCR	300

[illegible]

093296

Search completed: December 13, 2003, 13:29:48
Job time : 111.239 secs

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 26.3248 Seconds
(without alignments)
1375.531 Million cell updates/sec

Title: US-09-785-215-2
Perfect score: 4058
Sequence: 1 MLPGLALLLLAAWTARALEV.....KMQQNGYENPTYKFFEQMQN 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	4058	100.0	770	1	A4_HUMAN	P05067 h amyloid b
2	4040	99.6	770	1	A4_MACFA	P53601 m amyloid b
3	3981	98.1	770	1	A4_PIG	P79307 s amyloid b
4	3952	97.4	770	1	A4_CAVPO	Q60495 c amyloid b
5	3937	97.0	770	1	A4_RAT	P08592 r amyloid b
6	3923	96.7	770	1	A4_MOUSE	P12023 m amyloid b
7	3921.5	96.6	751	1	A4_SAISC	Q95241 s amyloid b
8	1981.5	48.8	763	1	APP2_HUMAN	Q06481 homo sapien
9	1971.5	48.6	765	1	APP2_RAT	P15943 rattus norv
10	1704.5	42.0	695	1	APP2_MOUSE	Q06335 mus musculu
11	1159.5	28.6	650	1	APP1_HUMAN	P51693 homo sapien
12	1150.5	28.4	653	1	APP1_MOUSE	Q03157 mus musculu
13	785	19.3	686	1	A4_CAEEL	Q10651 caenorhabdi
14	739	18.2	887	1	A4_DROME	P14599 drosophila
15	403	9.9	76	1	A4_MACMU	P29216 macaca mula
16	292	7.2	59	1	A4_BOVIN	Q28053 bos taurus
17	288	7.1	58	1	A4_RABIT	Q28748 oryctolagus

18	288	7.1	58	1	A4_SHEEP	Q28757	ovis aries
19	287	7.1	58	1	A4_CANFA	Q28280	canis famil
20	283	7.0	57	1	A4_URSMA	Q29149	ursus marit
21	194.5	4.8	3911	1	AKA9_HUMAN	Q99996	h a-kinase
22	186	4.6	252	1	SPT2_HUMAN	O43291	homo sapien
23	186	4.6	993	1	SCP1_MOUSE	Q62209	mus musculu
24	176	4.3	55	1	ISH1_STOHE	P31713	stoichactis
25	175.5	4.3	579	1	G160_HUMAN	Q08378	homo sapien
26	175	4.3	302	1	TFPI_RAT	Q02445	rattus norv
27	174.5	4.3	252	1	SPT2_MOUSE	Q9wu03	mus musculu
28	174.5	4.3	513	1	SPT1_HUMAN	O43278	homo sapien
29	169	4.2	1875	1	MLP1_YEAST	Q02455	saccharomyc
30	168	4.1	304	1	TFPI_HUMAN	P10646	homo sapien
31	166	4.1	55	1	ISH2_STOHE	P81129	stoichactis
32	164.5	4.1	507	1	SPT1_MOUSE	Q9r097	mus musculu
33	163.5	4.0	346	1	AMBP_MERUN	Q62577	meriones un
34	163.5	4.0	1130	1	YL17_CAEEL	Q11102	caenorhabdi
35	163.5	4.0	2130	1	BA2B_CHICK	Q9de13	gallus gall
36	163	4.0	2944	1	CA17_HUMAN	Q02388	homo sapien
37	163	4.0	3176	1	CA36_HUMAN	P12111	homo sapien
38	162.5	4.0	297	1	TRT2_HUMAN	P45379	homo sapien
39	162	4.0	304	1	TFPI_MACMU	Q28864	macaca mula
40	162	4.0	407	1	IE68_HSVSA	Q01042	herpesvirus
41	162	4.0	3137	1	CA36_CHICK	P15989	gallus gall
42	161.5	4.0	630	1	YCF2_OENVI	P31569	oenothera v
43	161.5	4.0	1278	1	DYNA_HUMAN	Q14203	homo sapien
44	161	4.0	64	1	SPT3_HUMAN	P49223	homo sapien
45	160.5	4.0	721	1	YCF2_OENPI	P31568	oenothera p

ALIGNMENTS

RESULT 1

A4_HUMAN

ID A4_HUMAN STANDARD; PRT; 770 AA.

AC P05067; P09000; P78438; Q13764; Q13778; Q13793; Q16011; Q9BT38;
AC Q9UCB6; Q9UQ58;

DT 13-AUG-1987 (Rel. 05, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE amyloid protein) (Cerebral vascular amyloid peptide) (CVAP) (Protease
DE nexin-II) (PN-II) (APPI) (PreA4) [Contains: Soluble APP-alpha (S-APP-
DE alpha); Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42
DE (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42);
DE P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59)
DE (Amyloid intracellular domain 59) (AID(59)); Gamma-CTF(57) (Gamma-
DE secretase C-terminal fragment 57) (Amyloid intracellular domain 57)
DE (AID(57)); Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50)
DE (Amyloid intracellular domain 50) (AID(50)); C31].

GN APP OR A4 OR AD1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM APP695).
 RC TISSUE=Brain;
 RX MEDLINE=87144572; PubMed=2881207;
 RA Kang J., Lemaire H.-G., Unterbeck A., Salbaum J.M., Masters C.L.,
 RA Grzeschik K.-H., Multhaup G., Beyreuther K., Mueller-Hill B.;
 RT "The precursor of Alzheimer's disease amyloid A4 protein resembles a
 RT cell-surface receptor.";
 RL Nature 325:733-736(1987).
 RN [2]

RP SEQUENCE FROM N.A. (ISOFORM APP751).
 RC TISSUE=Brain;
 RX MEDLINE=88122639; PubMed=2893289;
 RA Ponte P., Gonzalez-Dewhitt P., Schilling J., Miller J., Hsu D.,
 RA Greenberg B., Davis K., Wallace W., Lieberburg I., Fuller F.,
 RA Cordell B.;
 RT "A new A4 amyloid mRNA contains a domain homologous to serine
 RT proteinase inhibitors.";
 RL Nature 331:525-527(1988).
 RN [3]

RP SEQUENCE FROM N.A. (ISOFORM APP695).
 RX MEDLINE=89128427; PubMed=2783775;
 RA Lemaire H.-G., Salbaum J.M., Multhaup G., Kang J., Bayney R.M.,
 RA Unterbeck A., Beyreuther K., Mueller-Hill B.;
 RT "The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid
 RT is encoded by 16 exons.";
 RL Nucleic Acids Res. 17:517-522(1989).
 RN [4]

RP SEQUENCE FROM N.A. (ISOFORM APP770).
 RX MEDLINE=90236318; PubMed=2110105;
 RA Yoshikai S.-I., Sasaki H., Doh-Ura K., Furuya H., Sakaki Y.;
 RT "Genomic organization of the human amyloid beta-protein precursor
 RT gene.";
 RL Gene 87:257-263(1990).
 RN [5]

RP ERRATUM, AND REVISIONS.
 RA Yoshikai S.-I., Sasaki H., Doh-ura K., Furuya H., Sakaki Y.;
 RL Gene 102:291-292(1991).
 RN [6]

RP SEQUENCE FROM N.A. (ISOFORM L-APP733).
 RC TISSUE=Leukocyte;
 RX MEDLINE=92268136; PubMed=1587857;
 RA Koenig G., Moenning U., Czech C., Prior R., Banati R.,
 RA Schreiter-Gasser U., Bauer J., Masters C.L., Beyreuther K.;
 RT "Identification and differential expression of a novel alternative
 RT splice isoform of the beta A4 amyloid precursor protein (APP) mRNA in
 RT leukocytes and brain microglial cells.";
 RL J. Biol. Chem. 267:10804-10809(1992).
 RN [7]

RP SEQUENCE FROM N.A. (ISOFORM APP770).
 RX MEDLINE=97263807; PubMed=9108164;
 RA Hattori M., Tsukahara F., Furuhashi Y., Tanahashi H., Hirose M.,
 RA Saito M., Tsukuni S., Sakaki Y.;
 RT "A novel method for making nested deletions and its application for
 RT sequencing of a 300 kb region of human APP locus.";
 RL Nucleic Acids Res. 25:1802-1808(1997).
 RN [8]

RP SEQUENCE FROM N.A. (ISOFORM APP305).

RC TISSUE=Pancreas;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnierch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [9]
 RP SEQUENCE OF 1-10 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89016647; PubMed=3140222;
 RA Schon E.A., Mita S., Sadlock J., Herbert J.;
 RT "A cDNA specifying the human amyloid beta precursor protein (ABPP)
 RT encodes a 95-kDa polypeptide.";
 RL Nucleic Acids Res. 16:9351-9351(1988).
 RN [10]
 RP ERRATUM, AND REVISIONS.
 RA Mita S., Sadlock J., Herbert J., Schon E.A.;
 RL Nucleic Acids Res. 16:11402-11402(1988).
 RN [11]
 RP SEQUENCE OF 1-75 FROM N.A.
 RX MEDLINE=89165870; PubMed=2538123;
 RA La Fauci G., Lahiri D.K., Salton S.R., Robakis N.K.;
 RT "Characterization of the 5'-end region and the first two exons of the
 RT beta-protein precursor gene.";
 RL Biochem. Biophys. Res. Commun. 159:297-304(1989).
 RN [12]
 RP SEQUENCE OF 18-50.
 RC TISSUE=Fibroblast;
 RX MEDLINE=87250462; PubMed=3597385;
 RA van Nostrand W.E., Cunningham D.D.;
 RT "Purification of protease nexin II from human fibroblasts.";
 RL J. Biol. Chem. 262:8508-8514(1987).
 RN [13]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP751).
 RC TISSUE=Brain;
 RX MEDLINE=89346754; PubMed=2569763;
 RA de Sauvage F., Octave J.N.;
 RT "A novel mRNA of the A4 amyloid precursor gene coding for a possibly
 RT secreted protein.";
 RL Science 245:651-653(1989).
 RN [14]

RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP695).
 RC TISSUE=Brain;
 RX MEDLINE=87231971; PubMed=3035574;
 RA Robakis N.K., Ramakrishna N., Wolfe G., Wisniewski H.M.;
 RT "Molecular cloning and characterization of a cDNA encoding the
 RT cerebrovascular and the neuritic plaque amyloid peptides.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4190-4194(1987).
 RN [15]
 RP SEQUENCE OF 286-366 FROM N.A.
 RX MEDLINE=88122640; PubMed=2893290;
 RA Tanzi R.E., McClatchey A.I., Lamperti E.D., Villa-Komaroff L.,
 RA Gusella J.F., Neve R.L.;
 RT "Protease inhibitor domain encoded by an amyloid protein precursor
 RT mRNA associated with Alzheimer's disease.";
 RL Nature 331:528-530(1988).
 RN [16]
 RP SEQUENCE OF 287-367 FROM N.A.
 RX MEDLINE=88122641; PubMed=2893291;
 RA Kitaguchi N., Takahashi Y., Tokushima Y., Shiojiri S., Ito H.;
 RT "Novel precursor of Alzheimer's disease amyloid protein shows
 RT protease inhibitory activity.";
 RL Nature 331:530-532(1988).
 RN [17]
 RP SEQUENCE OF 507-770 FROM N.A.
 RC TISSUE=Brain cortex;
 RX MEDLINE=88124954; PubMed=2893379;
 RA Zain S.B., Salim M., Chou W.G., Sajdel-Sulkowska E.M., Majocha R.E.,
 RA Marotta C.A.;
 RT "Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer
 RT disease brain: coding and noncoding regions of the fetal precursor
 RT mRNA are expressed in the cortex.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:929-933(1988).
 RN [18]
 RP SEQUENCE OF 523-555, AND COLLAGEN-BINDING DOMAIN.
 RX MEDLINE=96139497; PubMed=8576160;
 RA Behr D., Hesse L., Masters C.L., Multhaup G.;
 RT "Regulation of amyloid protein precursor (APP) binding to collagen and
 RT mapping of the binding sites on APP and collagen type I.";
 RL J. Biol. Chem. 271:1613-1620(1996).
 RN [19]
 RP SEQUENCE OF 656-737 FROM N.A.
 RX MEDLINE=89392030; PubMed=2675837;
 RA Johnstone E.M., Chaney M.O., Moore R.E., Ward K.E., Norris F.H.,
 RA Little S.P.;
 RT "Alzheimer's disease amyloid peptide is encoded by two exons and shows
 RT similarity to soybean trypsin inhibitor.";
 RL Biochem. Biophys. Res. Commun. 163:1248-1255(1989).
 RN [20]
 RP SEQUENCE OF 672-681.
 RC TISSUE=Brain cortex;
 RX MEDLINE=88035004; PubMed=3312495;
 RA Pardridge W.M., Vinters H.V., Yang J., Eisenberg J., Choi T.B.,
 RA Tourtellotte W.W., Huebner V., Shively J.E.;
 RT "Amyloid angiopathy of Alzheimer's disease: amino acid composition
 RT and partial sequence of a 4,200-dalton peptide isolated from cortical
 RT microvessels.";
 RL J. Neurochem. 49:1394-1401(1987).

RN [21]
RP SEQUENCE OF 674-770 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=87120328; PubMed=3810169;
RA Goldgaber D., Lerman M.I., McBride O.W., Saffiotti U., Gajdusek D.C.;
RT "Characterization and chromosomal localization of a cDNA encoding
RT brain amyloid of Alzheimer's disease.";

Query Match 100.0%; Score 4058; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 6.6e-199;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDPSGK 60
      |||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDPSGK 60

Qy     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
      |||
Db     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

Qy    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181 GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE 240
      |||
Db    181 GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE 240

Qy    241 EADDDDEDEDGDEVEEEAEPPYEEATERTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241 EADDDDEDEDGDEVEEEAEPPYEEATERTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300

Qy    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420
      |||
Db    361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420

Qy    421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480
      |||
Db    421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480

Qy    481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTHLRVIYER 540
      |||
Db    481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTHLRVIYER 540

Qy    541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDAIMPSTTET 600
      |||
Db    541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDAIMPSTTET 600

Qy    601 KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 660
      |||
Db    601 KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Qy    661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
```

```

Db      661  IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAITVIVITL 720
QY      721  VMLKKKQYTSIHGHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db      721  VMLKKKQYTSIHGHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

```

RESULT 2

A4_MACFA

```

ID      A4_MACFA          STANDARD;          PRT;      770 AA.
AC      P53601; Q95KN7;
DT      01-OCT-1996 (Rel. 34, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE      amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);
DE      Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-
DE      APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40);
DE      Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)
DE      (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-
DE      secretase C-terminal fragment 50); C31].
GN      APP.
OS      Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC      Cercopithecinae; Macaca.
OX      NCBI_TaxID=9541;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORMS APP695 AND APP770).
RC      TISSUE=Cerebellum;
RX      MEDLINE=91273117; PubMed=1905108;
RA      Podlisny M.B., Tolan D.R., Selkoe D.J.;
RT      "Homology of the amyloid beta protein precursor in monkey and human
RT      supports a primate model for beta amyloidosis in Alzheimer's
RT      disease.";
RL      Am. J. Pathol. 138:1423-1435(1991).
CC      -!- FUNCTION: Functions as a cell surface receptor and performs
CC      physiological functions on the surface of neurons relevant to
CC      neurite growth, neuronal adhesion and axonogenesis. Involved in
CC      cell mobility and transcription regulation through protein-protein
CC      interactions (By similarity). Can promote transcription activation
CC      through binding to APBB1/Tip60 and inhibit Notch signaling through
CC      interaction with Numb (By similarity). Couples to apoptosis-
CC      inducing pathways such as those mediated by G(0) and JIP (By
CC      similarity). Inhibits G(0) alpha ATPase activity (By similarity).
CC      Acts as a kinesin I membrane receptor, mediating the axonal
CC      transport of beta-secretase and presenilin 1 (By similarity). May
CC      be involved in copper homeostasis/oxidative stress through copper
CC      ion reduction. In vitro, copper-metallated APP induces neuronal
CC      death directly or is potentiated through Cu(II)-mediated low-
CC      density lipoprotein oxidation (By similarity). Can regulate
CC      neurite outgrowth through binding to components of the
CC      extracellular matrix such as heparin and collagen I and IV (By
CC      similarity). The splice isoforms that contain the BPTI domain
CC      possess protease inhibitor activity (By similarity).
CC      -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators

```

CC with metal-reducing activity. Bind transient metals such as
 CC copper, zinc and iron (By similarity).

CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
 CC peptides, including C31, are potent enhancers of neuronal
 CC apoptosis (By similarity).

CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
 CC cytoplasmic proteins, including APBB family members, the APBA
 CC family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
 CC to Dab1 inhibits its serine phosphorylation (By similarity). Also
 CC interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
 CC (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
 CC In vitro, it binds MAPT via the MT-binding domains (By
 CC similarity). Associates with microtubules in the presence of ATP
 CC and in a kinesin-dependent manner (By similarity).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
 CC protein that rapidly becomes internalized via clatherin-coated
 CC pits. During maturation, the immature APP (N-glycosylated in the
 CC endoplasmic reticulum) moves to the Golgi complex where complete
 CC maturation occurs (O-glycosylated and sulfated). After alpha-
 CC secretase cleavage, soluble APP is released into the extracellular
 CC space and the C-terminal is internalized to endosomes and
 CC lysosomes. Some APP accumulates in secretory transport vesicles
 CC leaving the late Golgi compartment and returns to the cell
 CC surface. GammaCTF(59) peptide is located to both the cytoplasm and
 CC nuclei of neurons (By similarity).

CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms seem to exist;
 CC Name=APP770;
 CC IsoId=P53601-1; Sequence=Displayed;
 CC Name=APP695;
 CC IsoId=P53601-2; Sequence=VSP_000010, VSP_000011;

CC -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
 CC sorting of membrane proteins to the basolateral surface of
 CC epithelial cells (By similarity).

CC -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
 CC phosphorylated proteins is required for the specific binding of
 CC the PID domain. However additional amino acids either N- or C-
 CC terminal to the NPXY motif are often required for complete
 CC interaction. The PID domain-containing proteins which bind APP
 CC require the YENPTY motif for full interaction. These interactions
 CC are independent of phosphorylation on the terminal tyrosine
 CC residue. The NPXY site is also involved in clatherin-mediated
 CC endocytosis (By similarity).

CC -!- PTM: Proteolytically processed under normal cellular conditions.
 CC Cleavage by alpha-secretase or alternatively by beta-secretase
 CC leads to generation and extracellular release of soluble APP
 CC peptides, S-APP-alpha and S-APP-beta, respectively, and the
 CC retention of corresponding membrane-anchored C-terminal fragments,
 CC C83 and C99. Subsequent processing of C83 by gamma-secretase
 CC yields P3 peptides. This is the major secretory pathway and is
 CC nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated
 CC gamma-secretase processing of C99 releases the amyloid beta
 CC proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),
 CC major components of amyloid plaques, and the cytotoxic C-terminal
 CC fragments, gammaCTF(50), gammaCTF(57) and gammaCTF(59) (By
 CC similarity).

CC -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis
 CC (By similarity). Cleavage at Asp-739 by either caspase-3, -8 or -9
 CC results in the production of the neurotoxic C31 peptide and the
 CC increased production of beta-amyloid peptides (By similarity).
 CC -!- PTM: N- and O-linked glycosylated (By similarity).
 CC -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
 CC serine residues is neuron-specific. Phosphorylation can affect APP
 CC processing, neuronal differentiation and interaction with other
 CC proteins (By similarity).
 CC -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
 CC zinc, can induce histidine-bridging between beta-amyloid molecules
 CC resulting in beta-amyloid-metal aggregates (By similarity).
 CC Extracellular zinc-binding increases binding of heparin to APP and
 CC inhibits collagen-binding (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M58727; AAA36829.1; -.
 DR EMBL; M58726; AAA36828.1; -.
 DR HSSP; P05067; 1AAP.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.
 KW Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
 KW Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
 KW Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
 KW Proteoglycan; Alternative splicing; Amyloid.
 FT SIGNAL 1 17 BY SIMILARITY.
 FT CHAIN 18 770 AMYLOID BETA A4 PROTEIN.
 FT CHAIN 18 687 SOLUBLE APP-ALPHA (POTENTIAL).
 FT CHAIN 18 671 SOLUBLE APP-BETA (POTENTIAL).
 FT CHAIN 672 770 C99 (POTENTIAL).
 FT CHAIN 672 713 BETA-AMYLOID PROTEIN 42 (POTENTIAL).
 FT CHAIN 672 711 BETA-AMYLOID PROTEIN 40 (POTENTIAL).
 FT CHAIN 688 770 C83 (POTENTIAL).
 FT CHAIN 688 713 P3(42) (POTENTIAL).
 FT CHAIN 688 711 P3(40) (POTENTIAL).
 FT CHAIN 712 770 GAMMA-CTF(59) (POTENTIAL).
 FT CHAIN 714 770 GAMMA-CTF(57) (POTENTIAL).

FT	CHAIN	721	770	GAMMA-CTF(50) (POTENTIAL).
FT	CHAIN	740	770	C31 (POTENTIAL).
FT	DOMAIN	18	699	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	700	723	POTENTIAL.
FT	DOMAIN	724	770	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	96	110	HEPARIN-BINDING (BY SIMILARITY).
FT	DOMAIN	181	188	ZINC-BINDING (BY SIMILARITY).
FT	DOMAIN	291	341	BPTI/KUNITZ INHIBITOR.
FT	DOMAIN	391	423	HEPARIN-BINDING (BY SIMILARITY).
FT	DOMAIN	491	522	HEPARIN-BINDING (BY SIMILARITY).
FT	DOMAIN	523	540	COLLAGEN-BINDING (BY SIMILARITY).
FT	DOMAIN	732	751	INTERACTION WITH G(O)-ALPHA
FT				(BY SIMILARITY).
FT	DOMAIN	230	260	ASP/GLU-RICH (ACIDIC).
FT	DOMAIN	274	280	POLY-THR.
FT	SITE	144	144	REQUIRED FOR COPPER(II) REDUCTION
FT				(BY SIMILARITY).
FT	ACT_SITE	301	302	REACTIVE BOND (BY SIMILARITY).
FT	SITE	671	672	CLEAVAGE (BY BETA-SECRETASE)
FT				(BY SIMILARITY).
FT	SITE	672	673	CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT	SITE	687	688	CLEAVAGE (BY ALPHA-SECRETASE)
FT				(BY SIMILARITY).
FT	SITE	704	704	IMPLICATED IN FREE RADICAL PROPAGATION
FT				(BY SIMILARITY).
FT	SITE	706	706	INVOLVED IN OXIDATIVE REACTIONS
FT				(BY SIMILARITY).
FT	SITE	711	712	CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT				(BY SIMILARITY).
FT	SITE	713	714	CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT				(BY SIMILARITY).
FT	SITE	720	721	CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)
FT				(BY SIMILARITY).
FT	SITE	724	734	BASOLATERAL SORTING SIGNAL
FT				(BY SIMILARITY).
FT	SITE	739	740	CLEAVAGE (BY CASPASES-3,-6,-8 OR -9)
FT				(BY SIMILARITY).
FT	SITE	757	760	ENDOCYTOSIS SIGNAL.
FT	SITE	759	762	NPXY MOTIF.

Query Match 99.6%; Score 4040; DB 1; Length 770;
 Best Local Similarity 99.5%; Pred. No. 5.4e-198;
 Matches 766; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIYRCLVG	120
Db	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE	240

Db	181		GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Qy	241		EADDDDEDEDGDEVEEEAEEPYYEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Db	241		EADDDDEDEDGDEVEEEAEEPYYEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Qy	301		RAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301		RAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSVMSQSLRKTTRREPLTRD	360
Qy	361		PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Db	361		PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Qy	421		KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Db	421		KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Qy	481		QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481		QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541		MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRISYGNDALMPSLTET	600
Db	541		MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRISYGNDALMPSLTET	600
Qy	601		KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601		KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661		IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661		IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721		VMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770
Db	721		VMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770

RESULT 3

A4_PIG

ID A4_PIG STANDARD; PRT; 770 AA.

AC P79307; Q29023; Q9TUI0;

DT 01-NOV-1997 (Rel. 35, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha); Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50); C31].

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kimura A., Takahashi T.;
 RT "Amyloid precursor protein 770.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE OF 1-136 FROM N.A.
 RC TISSUE=Small intestine;
 RA Winteroe A.K., Fredholm M.;
 RT "Evaluation and characterization of a porcine small intestine cDNA
 RT library.";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE OF 667-723 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92017079; PubMed=1656157;
 RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
 RT "Conservation of the sequence of the Alzheimer's disease amyloid
 RT peptide in dog, polar bear and five other mammals by cross-species
 RT polymerase chain reaction analysis.";
 RL Brain Res. Mol. Brain Res. 10:299-305(1991).
 CC -!- FUNCTION: Functions as a cell surface receptor and performs
 CC physiological functions on the surface of neurons relevant to
 CC neurite growth, neuronal adhesion and axonogenesis. Involved in
 CC cell mobility and transcription regulation through protein-protein
 CC interactions (By similarity). Can promote transcription activation
 CC through binding to APBB1/Tip60 and inhibit Notch signaling through
 CC interaction with Numb (By similarity). Couples to apoptosis-
 CC inducing pathways such as those mediated by G(0) and JIP (By
 CC similarity). Inhibits G(0) alpha ATPase activity (By similarity).
 CC Acts as a kinesin I membrane receptor, mediating the axonal
 CC transport of beta-secretase and presenilin 1 (By similarity). May
 CC be involved in copper homeostasis/oxidative stress through copper
 CC ion reduction (By similarity). In vitro, copper-metallated APP
 CC induces neuronal death directly or is potentiated through Cu(II)-
 CC mediated low-density lipoprotein oxidation (By similarity). Can
 CC regulate neurite outgrowth through binding to components of the
 CC extracellular matrix such as heparin and collagen I and IV (By
 CC similarity).
 CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
 CC with metal-reducing activity. Bind transient metals such as
 CC copper, zinc and iron (By similarity).
 CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
 CC peptides, including C31, are potent enhancers of neuronal
 CC apoptosis (By similarity).
 CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
 CC cytoplasmic proteins, including APBB family members, the APBA
 CC family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
 CC to Dab1 inhibits its serine phosphorylation (By similarity). Also
 CC interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
 CC (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
 CC In vitro, it binds MAPT via the MT-binding domains (By
 CC similarity). Associates with microtubules in the presence of ATP
 CC and in a kinesin-dependent manner (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
 CC protein that rapidly becomes internalized via clatherin-coated

CC pits. During maturation, the immature APP (N-glycosylated in the
 CC endoplasmic reticulum) moves to the Golgi complex where complete
 CC maturation occurs (O-glycosylated and sulfated). After alpha-
 CC secretase cleavage, soluble APP is released into the extracellular
 CC space and the C-terminal is internalized to endosomes and
 CC lysosomes. Some APP accumulates in secretory transport vesicles
 CC leaving the late Golgi compartment and returns to the cell
 CC surface. GammaCTF(59) peptide is located to both the cytoplasm and
 CC nuclei of neurons (By similarity).

CC -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
 CC sorting of membrane proteins to the basolateral surface of
 CC epithelial cells (By similarity).

CC -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
 CC phosphorylated proteins is required for the specific binding of
 CC the PID domain. However additional amino acids either N- or C-
 CC terminal to the NPXY motif are often required for complete
 CC interaction. The PID domain-containing proteins which bind APP
 CC require the YENPTY motif for full interaction. These interactions
 CC are independent of phosphorylation on the terminal tyrosine
 CC residue. The NPXY site is also involved in clatherin-mediated
 CC endocytosis (By similarity).

CC -!- PTM: Proteolytically processed under normal cellular conditions.
 CC Cleavage by alpha-secretase or alternatively by beta-secretase
 CC leads to generation and extracellular release of soluble APP
 CC peptides, S-APP-alpha and S-APP-beta, respectively, and the
 CC retention of corresponding membrane-anchored C-terminal fragments,
 CC C83 and C99. Subsequent processing of C83 by gamma-secretase
 CC yields P3 peptides. This is the major secretory pathway and is
 CC nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated
 CC gamma-secretase processing of C99 releases the amyloid beta
 CC proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),
 CC major components of amyloid plaques, and the cytotoxic C-terminal
 CC fragments, gammaCTF(50), gammaCTF(57) and gammaCTF(59) (By
 CC similarity).

CC -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis
 CC (By similarity). Cleavage at Asp-739 by either caspase-3, -8 or -9
 CC results in the production of the neurotoxic C31 peptide and the
 CC increased production of beta-amyloid peptides (By similarity).

CC -!- PTM: N- and O-linked glycosylated (By similarity).

CC -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
 CC serine residues is neuron-specific. Phosphorylation can affect APP
 CC processing, neuronal differentiation and interaction with other
 CC proteins (By similarity).

CC -!- PTM: Extracellular binding and reduction of copper, results in a
 CC corresponding oxidation of Cys-144 and Cys-158, and the formation
 CC of a disulfide bond (By similarity).

CC -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
 CC zinc, can induce histidine-bridging between beta-amyloid molecules
 CC resulting in beta-amyloid-metal aggregates (By similarity).
 CC Extracellular zinc-binding increases binding of heparin to APP and
 CC inhibits collagen-binding (By similarity).

CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.

CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB032550; BAA84580.1; -.
 DR EMBL; Z84022; CAB06313.1; -.
 DR EMBL; X56127; CAA39592.1; -.
 DR HSSP; P05067; 1AAP.
 DR InterPro; IPR008155; A4_APP.
 DR InterPro; IPR008154; A4_extra.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.
 KW Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
 KW Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
 KW Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
 KW Amyloid.
 FT SIGNAL 1 17 BY SIMILARITY.
 FT CHAIN 18 770 AMYLOID BETA A4 PROTEIN.
 FT CHAIN 18 687 SOLUBLE APP-ALPHA (POTENTIAL).
 FT CHAIN 18 671 SOLUBLE APP-BETA (POTENTIAL).
 FT CHAIN 672 770 C99 (BY SIMILARITY).
 FT CHAIN 672 713 BETA-AMYLOID PROTEIN 42 (BY SIMILARITY).
 FT CHAIN 672 711 BETA-AMYLOID PROTEIN 40 (BY SIMILARITY).
 FT CHAIN 688 770 C83 (BY SIMILARITY).
 FT CHAIN 688 713 P3(42) (BY SIMILARITY).
 FT CHAIN 688 711 P3(40) (BY SIMILARITY).
 FT CHAIN 712 770 GAMMA-CTF(59).
 FT CHAIN 714 770 GAMMA-CTF(57).
 FT CHAIN 721 770 GAMMA-CTF(50) (BY SIMILARITY).
 FT CHAIN 740 770 C31 (DURING APOPTOSIS) (BY SIMILARITY).
 FT DOMAIN 18 699 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 700 723 POTENTIAL.
 FT DOMAIN 724 770 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 96 110 HEPARIN-BINDING (BY SIMILARITY).
 FT DOMAIN 135 155 COPPER-BINDING (BY SIMILARITY).
 FT DOMAIN 181 188 ZINC-BINDING (BY SIMILARITY).
 FT DOMAIN 291 341 BPTI/KUNITZ INHIBITOR.
 FT DOMAIN 391 423 HEPARIN-BINDING (BY SIMILARITY).
 FT DOMAIN 491 522 HEPARIN-BINDING (BY SIMILARITY).
 FT DOMAIN 523 540 COLLAGEN-BINDING (BY SIMILARITY).
 FT DOMAIN 732 751 INTERACTION WITH G(O)-ALPHA (BY
 FT SIMILARITY).
 FT DOMAIN 230 260 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 274 280 POLY-THR.
 FT SITE 144 144 REQUIRED FOR COPPER(II) REDUCTION
 FT (BY SIMILARITY).

FT	ACT_SITE	301	302	REACTIVE BOND (BY SIMILARITY).
FT	SITE	671	672	CLEAVAGE (BY BETA-SECRETASE)
FT				(BY SIMILARITY).
FT	SITE	672	673	CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT	SITE	687	688	CLEAVAGE (BY ALPHA-SECRETASE)
FT				(BY SIMILARITY).
FT	SITE	704	704	IMPLICATED IN FREE RADICAL PROPAGATION
FT				(BY SIMILARITY).
FT	SITE	706	706	INVOLVED IN OXIDATIVE REACTIONS
FT				(BY SIMILARITY).
FT	SITE	711	712	CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT				(BY SIMILARITY).
FT	SITE	713	714	CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT				(BY SIMILARITY).

Query Match 98.1%; Score 3981; DB 1; Length 770;
Best Local Similarity 97.8%; Pred. No. 5.3e-195;
Matches 753; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHMNVQNGKWDSDPSGTK	60
Db	1	MLPGLALVLLAAWTARALEVPTDGNAGLLAEPQVAMFCGKLNMHMNVQNGKWESDPSTGK	60
Qy	61	TCIDTKEGILQYCQEVPPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIYRCLVG	120
Db	61	TCIGTKEGILQYCQEVPPELQITNVVEANQPVTIQNWCKRSRKQCKTHTHIVIPIYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP CGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP CGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEDSDSVWGGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Db	181	GVEFVCCPLAEESDNIDSADAEEDSDSVWGGADTDYADGSEDKVVEVAEEEEVADV EEE	240
Qy	241	EADDDDEDGEDGEVEEEEAEPPYEATERTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Db	241	EAEDDEDDEDGEDGEVEEEEAEPPYEATERTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFYFDVTEGKCAPPFFYGGCGGNRRNNFDTEEYCMAVCGSAMSQSLLKTQEPLARD	360
Db	301	RAMI SRWFYFDVTEGKCAPPFFYGGCGGNRRNNFDTEEYCMAVCGSVMSQSLLKTQEHL PQD	360
Qy	361	PVKLPPTTAASTPDAVDKYLET PGDENEHAFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLET PGDENEHAFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQH TLKHFEHV RMVDPKKA AQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQH TLKHFEHV RMVDPKKA AQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNP VPAVAEEIQDEVD ELLQKEQNYSDDVL ANMISEPRI SYGNDALMP SILTET	600

Db 541 MNQSLSLLYNPVAVAEI IQDEVDELLQKEQNYSDVLANMISEPRI SYGNDALMPSLTET 600

Qy 601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 |||||

Db 601 KTTVELLPVNGEFSLDDLQPWHPFGVDSVPANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 |||||

Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720

Qy 721 VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 |||||

Db 721 VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 4

A4_CAVPO

ID A4_CAVPO STANDARD; PRT; 770 AA.

AC Q60495; Q60496;

DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
 DE amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);
 DE Soluble APP-beta (S-APP-beta); CTF-alpha; CTF-beta; Beta-amyloid
 DE protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); P3(42);
 DE P3(40); CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-
 DE CTF(57) (Gamma-secretase C-terminal fragment 57); C31].

GN APP.

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.

OX NCBI_TaxID=10141;

RN [1]

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RC TISSUE=Brain, and Liver;

RX MEDLINE=97236426; PubMed=9116031;

RA Beck M., Mueller D., Bigl V.;

RT "Amyloid precursor protein in Guinea pigs - complete cDNA sequence and
 RT alternative splicing.";

RL Biochim. Biophys. Acta 1351:17-21(1997).

RN [2]

RP INTERACTION OF BETA-APP40 WITH APOE.

RX MEDLINE=98007700; PubMed=9349544;

RA Martel C.L., Mackic J.B., Matsubara E., Governale S., Miguel C.,
 RA Miao W., McComb J.G., Frangione B., Ghiso J., Zlokovic B.V.;

RT "Isoform-specific effects of apolipoproteins E2, E3, and E4 on
 RT cerebral capillary sequestration and blood-brain barrier transport of
 RT circulating Alzheimer's amyloid beta.";

RL J. Neurochem. 69:1995-2004(1997).

RN [3]

RP PROCESSING.

RX MEDLINE=20084499; PubMed=10619481;

RA Beck M., Brueckner M.K., Holzer M., Kaap S., Pannicke T., Arendt T.,
 RA Bigl V.;

RT "Guinea-pig primary cell cultures provide a model to study expression
 RT and amyloidogenic processing of endogenous amyloid precursor

RT protein.";

RL Neuroscience 95:243-254(2000).

RN [4]

RP GAMMA-SECRETASE PROCESSING.

RX MEDLINE=20576391; PubMed=11035007;

RA Pinnix I., Musunuru U., Tun H., Sridharan A., Golde T., Eckman C.,

RA Ziani-Cherif C., Onstead L., Sambamurti K.;

RT "A novel gamma -secretase assay based on detection of the putative

RT C-terminal fragment-gamma of amyloid beta protein precursor.";

RL J. Biol. Chem. 276:481-487(2001).

CC -!- FUNCTION: Functions as a cell surface receptor and performs

CC physiological functions on the surface of neurons relevant to

CC neurite growth, neuronal adhesion and axonogenesis. Involved in

CC cell mobility and transcription regulation through protein-protein

CC interactions (By similarity). Can promote transcription activation

CC through binding to APBB1/Tip60 and inhibit Notch signaling through

CC interaction with Numb (By similarity). Couples to apoptosis-

CC inducing pathways such as those mediated by G(O) and JIP (By

CC similarity). Inhibits G(O) alpha ATPase activity (By similarity).

CC Acts as a kinesin I membrane receptor, mediating the axonal

CC transport of beta-secretase and presenilin 1 (By similarity). May

CC be involved in copper homeostasis/oxidative stress through copper

CC ion reduction (By similarity). In vitro, copper-metallated APP

CC induces neuronal death directly or is potentiated through Cu(II)-

CC mediated low-density lipoprotein oxidation (By similarity). Can

CC regulate neurite outgrowth through binding to components of the

CC extracellular matrix such as heparin and collagen I and IV (By

CC similarity). The splice isoforms that contain the BPTI domain

CC possess protease inhibitor activity (By similarity).

CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators

CC with metal-reducing activity. Bind transient metals such as

CC copper, zinc and iron. Beta-amyloid peptides bind to lipoproteins

CC and apolipoproteins E and J in the CSF and to HDL particles in

CC plasma, inhibiting metal-catalyzed oxidation of lipoproteins.

CC -!- FUNCTION: Appicans elicit adhesion of neural cells to the

CC extracellular matrix and may regulate neurite outgrowth in the

CC brain (By similarity).

CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved

CC peptides, including C31, are potent enhancers of neuronal

CC apoptosis (By similarity).

CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several

CC cytoplasmic proteins, including APBB family members, the APBA

CC family, MAPK8IP1, SHC1 and Numb and Dab1 (By similarity). Also

CC interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2

CC (via its TPR domains), APPBP2 (via BaSS) and DDB1 (By similarity).

CC Associates with microtubules in the presence of ATP and in a

CC kinesin-dependent manner (By similarity). Soluble Abeta40 binds

CC all three isoforms of APOE, in vitro and in vivo. When lipidated,

CC ApoE3 appears to be the preferred amyloid binding isoform, while

CC the apoE4 isoform-beta-APP40 complex is capable of being

CC transported across the blood-brain barrier.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface

CC protein that rapidly becomes internalized via clathrin-coated pits

CC (By similarity). During maturation, the immature APP (N-

CC glycosylated in the endoplasmic reticulum) moves to the Golgi

CC complex where complete maturation occurs (O-glycosylated and

CC sulfated) (By similarity). After alpha-secretase cleavage, soluble

CC APP is released into the extracellular space and the C-terminal is
 CC internalized to endosomes and lysosomes (By similarity). Some APP
 CC accumulates in secretory transport vesicles leaving the late Golgi
 CC compartment and returns to the cell surface (By similarity). APP
 CC sorts to the basolateral surface in epithelial cells (By
 CC similatity).

CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms, missing exons 7,8 and 15, seem to
 CC exist. The L-isoforms, missing exon 15, are referred to as
 CC appicans;
 CC Name=APP770;
 CC IsoId=Q60495-1; Sequence=Displayed;
 CC Name=APP695;
 CC IsoId=Q60495-2; Sequence=VSP_007221, VSP_007222;

CC -!- TISSUE SPECIFICITY: Isoform APP695 is the major isoform found in
 CC brain. The longer isoforms containing the BPTI domain are
 CC predominantly expressed in peripheral organs such as muscle and
 CC liver.

CC -!- INDUCTION: Increased levels during neuronal differentiation.

CC -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
 CC sorting of membrane proteins to the basolateral surface of
 CC epithelial cells.

CC -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
 CC phosphorylated proteins is required for the specific binding of
 CC the PID domain. However additional amino acids either N- or C-
 CC terminal to the NPXY motif are often required for complete
 CC interaction. The PID domain-containing proteins which bind APP
 CC require the YENPTY motif for full interaction. These interactions
 CC are independent of phosphorylation on the terminal tyrosine
 CC residue (By similarity). The NPXY site is also involved in
 CC clatherin-mediated endocytosis.

CC -!- PTM: Proteolytically processed under normal cellular conditions.
 CC Cleavage by alpha-secretase or alternatively by beta-secretase
 CC leads to generation and extracellular release of soluble APP
 CC peptides, S-APP-alpha and S-APP-beta, respectively, and the
 CC retention of corresponding membrane-anchored C-terminal fragments,
 CC CTF-alpha and CTF-beta. Subsequent processing of CTF-alpha by
 CC gamma-secretase yields P3 peptides. This is the major secretory
 CC pathway and is nonamyloidogenic. Alternatively,
 CC presenilin/nicastrin-mediated gamma-secretase processing of CTF-
 CC beta releases the amyloid beta proteins, amyloid-beta 40 (Abeta40)
 CC and amyloid-beta 42 (Abeta42), major components of amyloid
 CC plaques, and the corresponding cytotoxic C-terminal fragments
 CC (CTFs).

CC -!- PTM: Proteolytically cleaved by caspase-3 during neuronal
 CC apoptosis (By similarity).

CC -!- PTM: N- and O-linked glycosylated. O-linkage of chondroitin
 CC sulfate to the L-APP isoforms produces the APP proteoglycan core
 CC proteins, the appicans (By similarity).

CC -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
 CC serine residues is neuron-specific (By similarity).
 CC Phosphorylation can affect APP processing, neuronal
 CC differentiation and interaction with other proteins.

CC -!- PTM: Extracellular binding and reduction of copper, results in a
 CC corresponding oxidation of Cys-144 and Cys-158, and the formation
 CC of a disulfide bond (By similarity).

DE amyloid protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40);
DE C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal
DE fragment 59); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57);
DE Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50); C31].
GN APP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM APP695).
RC TISSUE=Brain;
RX MEDLINE=88312583; PubMed=2900758;
RA Shivers B.D., Hilbich C., Multhaup G., Salbaum J.M., Beyreuther K.,
RA Seeburg P.H.;
RT "Alzheimer's disease amyloidogenic glycoprotein: expression pattern
RT in rat brain suggests a role in cell contact.";
RL EMBO J. 7:1365-1370(1988).
RN [2]
RP SEQUENCE OF 289-364 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89183625; PubMed=2648331;
RA Kang J., Mueller-Hill B.;
RT "The sequence of the two extra exons in rat preA4.";
RL Nucleic Acids Res. 17:2130-2130(1989).
RN [3]
RP SEQUENCE OF 720-730, AND MASS SPECTROMETRY.
RX PubMed=11483588;
RA Gu Y., Misonou H., Sato T., Dohmae N., Takio K., Ihara Y.;
RT "Distinct intramembrane cleavage of the beta-amyloid precursor protein
RT family resembling gamma-secretase-like cleavage of Notch.";
RL J. Biol. Chem. 276:35235-35238(2001).
RN [4]
RP ALTERNATIVE SPLICING.
RX PubMed=8624099;
RA Sandbrink R., Masters C.L., Beyreuther K.;
RT "APP gene family. Alternative splicing generates functionally related
RT isoforms.";
RL Ann. N.Y. Acad. Sci. 777:281-287(1996).
RN [5]
RP TISSUE SPECIFICITY OF APPICAN.
RX PubMed=7744833;
RA Shioi J., Pangalos M.N., Ripellino J.A., Vassilacopoulou D.,
RA Mytilineou C., Margolis R.U., Robakis N.K.;
RT "The Alzheimer amyloid precursor proteoglycan (appican) is present in
RT brain and is produced by astrocytes but not by neurons in primary
RT neural cultures.";
RL J. Biol. Chem. 270:11839-11844(1995).
RN [6]
RP TISSUE SPECIFICITY OF ISOFORMS.
RX PubMed=8996834;
RA Sandbrink R., Monning U., Masters C.L., Beyreuther K.;
RT "Expression of the APP gene family in brain cells, brain development
RT and aging.";
RL Gerontology 43:119-131(1997).
RN [7]
RP INTERACTION WITH DDB1, AND MUTAGENESIS OF TYR-757; ASN-759 AND

RP TYR-762.
 RX PubMed=9930726;
 RA Watanabe T., Sukegawa J., Tomita S., Iijima K.-I., Oguchi S.,
 RA Suzuki T., Nairn A.C., Greengard P.;
 RT "A 127-kDa protein (UV-DDB) binds to the cytoplasmic domain of the
 RT Alzheimer's amyloid precursor protein.";
 RL J. Neurochem. 72:549-556(1999).
 RN [8]
 RP INTERACTION WITH GNAO1, AND MUTAGENESIS OF HIS-732 AND HIS-733.
 RX PubMed=10024358;
 RA Brouillet E., Trembleau A., Galanaud D., Volovitch M., Bouillot C.,
 RA Valenza C., Prochiantz A., Allinquant B.;
 RT "The amyloid precursor protein interacts with Go heterotrimeric
 RT protein within a cell compartment specialized in signal
 RT transduction.";
 RL J. Neurosci. 19:1717-1727(1999).
 RN [9]
 RP CHARACTERISTICS OF APPICAN, AND MUTAGENESIS OF SER-656.
 RX MEDLINE=95256193; PubMed=7737970;
 RA Pangalos M.N., Efthimiopoulos S., Shioi J., Robakis N.K.;
 RT "The chondroitin sulfate attachment site of appican is formed by
 RT splicing out exon 15 of the amyloid precursor gene.";
 RL J. Biol. Chem. 270:10388-10391(1995).
 RN [10]
 RP BETA-AMYLOID METAL-BINDING.
 RX PubMed=10386999;
 RA Huang X., Atwood C.S., Hartshorn M.A., Multhaup G., Goldstein L.E.,
 RA Scarpa R.C., Cuajungco M.P., Gray D.N., Lim J., Moir R.D., Tanzi R.E.,
 RA Bush A.I.;
 RT "The A beta peptide of Alzheimer's disease directly produces hydrogen
 RT peroxide through metal ion reduction.";
 RL Biochemistry 38:7609-7616(1999).
 RN [11]
 RP BETA-AMYLOID ZINC BINDING.
 RX MEDLINE=99343552; PubMed=10413512;
 RA Liu S.T., Howlett G., Barrow C.J.;
 RT "Histidine-13 is a crucial residue in the zinc ion-induced aggregation
 RT of the A beta peptide of Alzheimer's disease.";
 RL Biochemistry 38:9373-9378(1999).
 RN [12]
 RP IMPORTANCE OF GLY-704 IN FREE RADICAL PROPAGATION, AND MUTAGENESIS OF
 RP GLY-704.
 RX PubMed=11959460;
 RA Kanski J., Varadarajan S., Aksenova M., Butterfield D.A.;
 RT "Role of glycine-33 and methionine-35 in Alzheimer's amyloid beta-
 RT peptide 1-42-associated oxidative stress and neurotoxicity.";
 RL Biochim. Biophys. Acta 1586:190-198(2001).
 RN [13]
 RP PHOSPHORYLATION.
 RX PubMed=9085254;
 RA Oishi M., Nairn A.C., Czernik A.J., Lim G.S., Isohara T., Gandy S.E.,
 RA Greengard P., Suzuki T.;
 RT "The cytoplasmic domain of Alzheimer's amyloid precursor protein is
 RT phosphorylated at Thr654, Ser655, and Thr668 in adult rat brain and
 RT cultured cells.";
 RL Mol. Med. 3:111-123(1997).
 RN [14]

RP PHOSPHORYLATION ON SER-730.
RX PubMed=10329382;
RA Isohara T., Horiuchi A., Watanabe T., Ando K., Czernik A.J., Uno I.,
RA Greengard P., Nairn A.C., Suzuki T.;
RT "Phosphorylation of the cytoplasmic domain of Alzheimer's beta-amyloid
RT precursor protein at Ser655 by a novel protein kinase.";
RL Biochem. Biophys. Res. Commun. 258:300-305(1999).
RN [15]
RP PHOSPHORYLATION, INDUCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
RP THR-743.
RX MEDLINE=99274744; PubMed=10341243;
RA Ando K., Oishi M., Takeda S., Iijima K.-I., Isohara T., Nairn A.C.,
RA Kirino Y., Greengard P., Suzuki T.;
RT "Role of phosphorylation of Alzheimer's amyloid precursor protein
RT during neuronal differentiation.";
RL J. Neurosci. 19:4421-4427(1999).
RN [16]
RP PHOSPHORYLATION ON THR-743.
RX PubMed=10936190;
RA Iijima K.-I., Ando K., Takeda S., Satoh Y., Seki T., Itohara S.,
RA Greengard P., Kirino Y., Nairn A.C., Suzuki T.;
RT "Neuron-specific phosphorylation of Alzheimer's beta-amyloid precursor
RT protein by cyclin-dependent kinase 5.";
RL J. Neurochem. 75:1085-1091(2000).
RN [17]
RP CARBOHYDRATE STRUCTURE OF APPICAN.
RX PubMed=11479316;
RA Tsuchida K., Shioi J., Yamada S., Boghosian G., Wu A., Cai H.,
RA Sugahara K., Robakis N.K.;
RT "Appican, the proteoglycan form of the amyloid precursor protein,
RT contains chondroitin sulfate E in the repeating disaccharide region
RT and 4-O-sulfated galactose in the linkage region.";
RL J. Biol. Chem. 276:37155-37160(2001).
CC -!- FUNCTION: Functions as a cell surface receptor and performs
CC physiological functions on the surface of neurons relevant to
CC neurite growth, neuronal adhesion and axonogenesis. Involved in
CC cell mobility and transcription regulation through protein-protein
CC interactions (By similarity). Can promote transcription activation
CC through binding to APBB1/Tip60 and inhibit Notch signaling through
CC interaction with Numb (By similarity). Couples to apoptosis-
CC inducing pathways such as those mediated by G(0) and JIP. Inhibits
CC G(0) alpha ATPase activity. Acts as a kinesin I membrane receptor,
CC mediating the axonal transport of beta-secretase and presenilin 1
CC (By similarity). May be involved in copper homeostasis/oxidative
CC stress through copper ion reduction. Can regulate neurite
CC outgrowth through binding to components of the extracellular
CC matrix such as heparin and collagen I and IV (By similarity). The
CC splice isoforms that contain the BPTI domain possess protease
CC inhibitor activity (By similarity).
CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC with metal-reducing activity. Bind transient metals such as
CC copper, zinc and iron. Rat and mouse beta-amyloid peptides bind
CC only weakly transient metals and have little reducing activity due
CC to substitutions of transient metal chelating residues. Beta-APP42
CC may activate mononuclear phagocytes in the brain and elicit
CC inflammatory responses. Promotes both tau aggregation and TPK II-
CC mediated phosphorylation (By similarity).

[illegible]

A4 MOUSE

```

AC      P12023; P97487; P97942; Q99K32;
DT      01-OCT-1989 (Rel. 12, Created)
DT      15-SEP-2003 (Rel. 42, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE      amyloid protein homolog) (Amyloidogenic glycoprotein) (AG) [Contains:
DE      Soluble APP-alpha (S-APP-alpha); Soluble APP-beta (S-APP-beta); C99
DE      (APP-C99); Beta-amyloid protein 42 (Beta-APP42); Beta-amyloid protein
DE      40 (Beta-APP40); C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase
DE      C-terminal fragment 59) (Amyloid intracellular domain 59) (AID(59))
DE      (APP-C59); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57)
DE      (Amyloid intracellular domain 57) (AID(57)) (APP-C57); Gamma-CTF(50)
DE      (Gamma-secretase C-terminal fragment 50) (Amyloid intracellular domain
DE      50) (AID(50)); C31].
GN      APP.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORM APP695).
RC      TISSUE=Brain;
RX      MEDLINE=88106489; PubMed=3322280;
RA      Yamada T., Sasaki H., Furuya H., Miyata T., Goto I., Sakaki Y.;
RT      "Complementary DNA for the mouse homolog of the human amyloid beta
RT      protein precursor.";
RL      Biochem. Biophys. Res. Commun. 149:665-671(1987).
RN      [2]
RP      REVISIONS.
RA      Yamada T.;
RL      Submitted (MAR-1988) to the EMBL/GenBank/DDBJ databases.
RN      [3]

```

RP SEQUENCE FROM N.A. (ISOFORM APP695).
 RC STRAIN=BALB/c; TISSUE=Brain;
 RX MEDLINE=92096458; PubMed=1756177;
 RA de Strooper B., van Leuven F., van den Berghe H.;
 RT "The amyloid beta protein precursor or proteinase nexin II from mouse
 RT is closer related to its human homolog than previously reported.";
 RL Biochim. Biophys. Acta 1129:141-143(1991).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM APP695).
 RC STRAIN=SAMP8; TISSUE=Hippocampus;
 RX PubMed=11235921;
 RA Kumar V.B., Vyas K., Franko M., Choudhary V., Buddhiraju C.,
 RA Alvarez J., Morley J.E.;
 RT "Molecular cloning, expression, and regulation of hippocampal amyloid
 RT precursor protein of senescence accelerated mouse (SAMP8).";
 RL Biochem. Cell Biol. 79:57-67(2001).
 RN [5]
 RP SEQUENCE OF 1-19 FROM N.A.
 RX MEDLINE=92209998; PubMed=1555768;
 RA Izumi R., Yamada T., Yoshikai S.I., Sasaki H., Hattori M.,
 RA Sakai Y.;
 RT "Positive and negative regulatory elements for the expression of the
 RT Alzheimer's disease amyloid precursor-encoding gene in mouse.";
 RL Gene 112:189-195(1992).
 RN [6]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP770).
 RC TISSUE=Breast tumor;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP SEQUENCE OF 281-380 FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Brain, and Kidney;
 RX MEDLINE=89149813; PubMed=2493250;
 RA Yamada T., Sasaki H., Dohura K., Goto I., Sakaki Y.;
 RT "Structure and expression of the alternatively-spliced forms of mRNA
 RT for the mouse homolog of Alzheimer's disease amyloid beta protein
 RT precursor.";
 RL Biochem. Biophys. Res. Commun. 158:906-912(1989).

RN [8]
 RP SEQUENCE OF 289-364 FROM N.A.
 RC STRAIN=CD-1; TISSUE=Placenta;
 RX MEDLINE=89345111; PubMed=2569710;
 RA Fukuchi K., Martin G.M., Deeb S.S.;
 RT "Sequence of the protease inhibitor domain of the A4 amyloid protein
 RT precursor of *Mus domesticus*.";
 RL Nucleic Acids Res. 17:5396-5396(1989).
 RN [9]
 RP SEQUENCE OF 656-737 FROM N.A.
 RC STRAIN=129/Sv;
 RA Wragg M.A., Busfield F., Duff K., Korenblat K., Capecchi M.,
 RA Loring J.F., Goate A.M.;
 RT "Introduction of six mutations into the mouse genome using 'Hit and
 RT Run' gene-targeting: introduction of familial Alzheimer's disease
 RT mutations into the mouse amyloid precursor protein gene and
 RT humanization of the A-beta fragment.";
 RL Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
 RN [10]
 RP TISSUE SPECIFICITY OF ALTERNATIVE SPLICED FORMS.
 RX PubMed=8510506;
 RA Sola C., Mengod G., Ghetti B., Palacios J.M., Triarhou L.C.;
 RT "Regional distribution of the alternatively spliced isoforms of beta
 RT APP RNA transcript in the brain of normal, heterozygous and
 RT homozygous weaver mutant mice as revealed by in situ hybridization
 RT histochemistry.";
 RL Brain Res. Mol. Brain Res. 17:340-346(1993).
 RN [11]
 RP INTERACTION WITH KNS2.
 RX PubMed=11144355;
 RA Kamal A., Stokin G.B., Yang Z., Xia C.-H., Goldstein L.S.;
 RT "Axonal transport of amyloid precursor protein is mediated by direct
 RT binding to the kinesin light chain subunit of kinesin-I.";
 RL Neuron 28:449-459(2000).
 RN [12]
 RP C-TERMINAL PROTEIN-PROTEIN INTERACTIONS, AND MUTAGENESIS OF TYR-728;
 RP THR-743; TYR-757; ASN-759 AND TYR-762.
 RX MEDLINE=21408156; PubMed=11517249;
 RA Matsuda S., Yasukawa T., Homma Y., Ito Y., Niikura T., Hiraki T.,
 RA Hirai S., Ohno S., Kita Y., Kawasumi M., Kouyama K., Yamamoto T.,
 RA Kyriakis J.M., Nishimoto I.;
 RT "C-jun N-terminal kinase (JNK)-interacting protein-1b/islet-brain-1
 RT scaffolds Alzheimer's amyloid precursor protein with JNK.";
 RL J. Neurosci. 21:6597-6607(2001).
 RN [13]
 RP INTERACTION WITH MAPK8IP1, AND PHOSPHORYLATION.
 RX MEDLINE=22028091; PubMed=11912189;
 RA Taru H., Iijima K.-I., Hase M., Kirino Y., Yagi Y., Suzuki T.;
 RT "Interaction of Alzheimer's beta-amyloid precursor family proteins
 RT with scaffold proteins of the JNK signaling cascade.";
 RL J. Biol. Chem. 277:20070-20078(2002).
 RN [14]
 RP INTERACTION OF CTF PEPTIDES WITH NUMB.
 RX PubMed=12011466;
 RA Roncarati R., Sestan N., Scheinfeld M.H., Berechid B.E., Lopez P.A.,
 RA Meucci O., McGlade J.C., Rakic P., D'Adamio L.;
 RT "The gamma-secretase-generated intracellular domain of beta-amyloid

RT precursor protein binds Numb and inhibits Notch signaling.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:7102-7107(2002).
 RN [15]
 RP GAMMA-SECRETASE PROCESSING, AND INTERACTION WITH APBB1.
 RX PubMed=11553691;
 RA Cupers P., Orlans I., Craessaerts K., Annaert W., De Strooper B.;
 RT "The amyloid precursor protein (APP)-cytoplasmic fragment generated by
 RT gamma-secretase is rapidly degraded but distributes partially in a
 RT nuclear fraction of neurones in culture.";
 RL J. Neurochem. 78:1168-1178(2001).
 CC -!- FUNCTION: Functions as a cell surface receptor and performs
 CC physiological functions on the surface of neurons relevant to
 CC neurite growth, neuronal adhesion and axonogenesis. Involved in
 CC cell mobility and transcription regulation through protein-protein
 CC interactions. Can promote transcription activation through binding
 CC to APBB1/Tip60 and inhibit Notch signaling through interaction
 CC with Numb. Couples to apoptosis-inducing pathways such as those
 CC mediated by G(O) and JIP. Inhibits G(O) alpha ATPase activity (By
 CC similarity). Acts as a kinesin I membrane receptor, mediating the
 CC axonal transport of beta-secretase and presenilin 1. May be
 CC involved in copper homeostasis/oxidative stress through copper ion
 CC reduction. Can regulate neurite outgrowth through binding to
 CC components of the extracellular matrix such as heparin and
 CC collagen I and IV (By similarity). The splice isoforms that
 CC contain the BPTI domain possess protease inhibitor activity (By
 CC similarity).
 CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
 CC with metal-reducing activity. Bind transient metals such as
 CC copper, zinc and iron. Rat and mouse beta-amyloid peptides bind
 CC only weakly transient metals and have little reducing activity due
 CC to substitutions of transient metal chelating residues. Beta-APP42
 CC may activate mononuclear phagocytes in the brain and elicit
 CC inflammatory responses. Promotes both tau aggregation and TPK II-
 CC mediated phosphorylation (By similarity).
 CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
 CC peptides, including C31, are potent enhancers of neuronal
 CC apoptosis.
 CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
 CC cytoplasmic proteins, including APBB family members, the APBA
 CC family, MAPK8IP1, SHC1, Numb and Dab1. Binding to Dab1 inhibits
 CC its serine phosphorylation. Also interacts with GPCR-like protein
 CC BPP, FPRL1, APPBP1, IB1, KNS2 (via its TPR domains), APPBP2 (via
 CC BaSS) and DDB1 (By similarity). In vitro, it binds MAPT via the
 CC MT-binding domains (By similarity). Associates with microtubules
 CC in the presence of ATP and in a kinesin-dependent manner (By
 CC similarity). Interacts, through a C-terminal domain, with GNAO1
 CC (By similarity). Amyloid beta-42 binds CHRNA7 in hippocampal
 CC neurons (By similarity). Beta-amyloid associates with HADH2 (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
 CC protein that rapidly becomes internalized via clatherin-coated
 CC pits. During maturation, the immature APP (N-glycosylated in the
 CC endoplasmic reticulum) moves to the Golgi complex where complete

Query Match 96.7%; Score 3923; DB 1; Length 770;
 Best Local Similarity 96.6%; Pred. No. 4.7e-192;
 Matches 744; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

AC Q95241;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Amyloid beta A4 protein precursor (APP) (Alzheimer's disease amyloid
 DE protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha); Soluble
 DE APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-APP42);
 DE Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40); Gamma-
 DE CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)
 DE (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-
 DE secretase C-terminal fragment 50); C31].
 GN APP.
 OS Saimiri sciureus (Common squirrel monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
 OX NCBI_TaxID=9521;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney, and Liver;
 RX MEDLINE=96108492; PubMed=8532114;
 RA Levy E., Amorim A., Frangione B., Walker L.C.;
 RT "Beta-amyloid precursor protein gene in squirrel monkeys with
 RT cerebral amyopathy."
 RL Neurobiol. Aging 16:805-808(1995).
 CC -!- FUNCTION: Functions as a cell surface receptor and performs
 CC physiological functions on the surface of neurons relevant to
 CC neurite growth, neuronal adhesion and axonogenesis. Involved in
 CC cell mobility and transcription regulation through protein-protein
 CC interactions (By similarity). Can promote transcription activation
 CC through binding to APBB1/Tip60 and inhibit Notch signaling through
 CC interaction with Numb (By similarity). Couples to apoptosis-
 CC inducing pathways such as those mediated by G(O) and JIP (By
 CC similarity). Inhibits G(O) alpha ATPase activity (By similarity).
 CC Acts as a kinesin I membrane receptor, mediating the axonal
 CC transport of beta-secretase and presenilin 1 (By similarity). May
 CC be involved in copper homeostasis/oxidative stress through copper
 CC ion reduction. In vitro, copper-metallated APP induces neuronal
 CC death directly or is potentiated through Cu(II)-mediated low-
 CC density lipoprotein oxidation (By similarity). Can regulate
 CC neurite outgrowth through binding to components of the
 CC extracellular matrix such as heparin and collagen I and IV (By
 CC similarity). The splice isoforms that contain the BPTI domain
 CC possess protease inhibitor activity (By similarity).
 CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
 CC with metal-reducing activity. Bind transient metals such as
 CC copper, zinc and iron (By similarity).
 CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
 CC peptides, including C31, are potent enhancers of neuronal
 CC apoptosis (By similarity).
 CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
 CC cytoplasmic proteins, including APBB family members, the APBA
 CC family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
 CC to Dab1 inhibits its serine phosphorylation (By similarity). Also
 CC interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
 CC (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
 CC In vitro, it binds MAPT via the MT-binding domains (By
 CC similarity). Associates with microtubules in the presence of ATP

CC and in a kinesin-dependent manner (By similarity).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface

CC protein that rapidly becomes internalized via clatherin-coated

CC pits. During maturation, the immature APP (N-glycosylated in the

CC endoplasmic reticulum) moves to the Golgi complex where complete

CC maturation occurs (O-glycosylated and sulfated). After alpha-

CC secretase cleavage, soluble APP is released into the extracellular

CC space and the C-terminal is internalized to endosomes and

CC lysosomes. Some APP accumulates in secretory transport vesicles

CC leaving the late Golgi compartment and returns to the cell

CC surface. GammaCTF(59) peptide is located to both the cytoplasm and

CC nuclei of neurons (By similarity).

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Comment=Additional isoforms seem to exist;

CC Name=APP770;

CC IsoId=Q95241-1; Sequence=Displayed;

CC Name=APP695;

CC IsoId=Q95241-2; Sequence=Not described;

CC -!- DOMAIN: The basolateral sorting signal (BaSS) is required for

CC sorting of membrane proteins to the basolateral surface of

CC epithelial cells (By similarity).

CC -!- DOMAIN: The NPXY sequence motif found in many tyrosine-

CC phosphorylated proteins is required for the specific binding of

CC the PID domain. However additional amino acids either N- or C-

CC terminal to the NPXY motif are often required for complete

CC interaction. The PID domain-containing proteins which bind APP

CC require the YENPTY motif for full interaction. These interactions

CC are independent of phosphorylation on the terminal tyrosine

CC residue. The NPXY site is also involved in clatherin-mediated

CC endocytosis (By similarity).

CC -!- PTM: Proteolytically processed under normal cellular conditions.

CC Cleavage by alpha-secretase or alternatively by beta-secretase

CC leads to generation and extracellular release of soluble APP

CC peptides, S-APP-alpha and S-APP-beta, respectively, and the

CC retention of corresponding membrane-anchored C-terminal fragments,

CC C83 and C99. Subsequent processing of C83 by gamma-secretase

CC yields P3 peptides. This is the major secretory pathway and is

CC nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated

CC gamma-secretase processing of C99 releases the amyloid beta

CC proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),

CC major components of amyloid plaques, and the cytotoxic C-terminal

CC fragments, gammaCTF(50), gammaCTF(57) and gammaCTF(59) (By

CC similarity).

CC -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis

CC (By similarity). Cleavage at Asp-720 by either caspase-3, -8 or -9

CC results in the production of the neurotoxic C31 peptide and the

CC increased production of beta-amyloid peptides (By similarity).

CC -!- PTM: N- and O-linked glycosylated (By similarity).

CC -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and

CC serine residues is neuron-specific. Phosphorylation can affect APP

CC processing, neuronal differentiation and interaction with other

CC proteins (By similarity).

CC -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and

CC zinc, can induce histidine-bridging between beta-amyloid molecules

CC resulting in beta-amyloid-metal aggregates (By similarity).

CC Extracellular zinc-binding increases binding of heparin to APP and

CC inhibits collagen-binding (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S81024; AAD14347.1; -.
 DR HSSP; P05067; 1AAP.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.
 KW Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
 KW Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
 KW Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
 KW Proteoglycan; Amyloid; Alternative splicing.
 FT SIGNAL 1 17 BY SIMILARITY.
 FT CHAIN 18 751 A4 PROTEIN.
 FT CHAIN 18 668 SOLUBLE APP-ALPHA (POTENTIAL).
 FT CHAIN 18 652 SOLUBLE APP-BETA (POTENTIAL).
 FT CHAIN 653 751 C99 (POTENTIAL).
 FT CHAIN 653 694 BETA-AMYLOID PROTEIN 42 (POTENTIAL).
 FT CHAIN 653 692 BETA-AMYLOID PROTEIN 40 (POTENTIAL).
 FT CHAIN 669 751 C83 (POTENTIAL).
 FT CHAIN 669 694 P3(42) (POTENTIAL).
 FT CHAIN 669 692 P3(40) (POTENTIAL).
 FT CHAIN 693 751 GAMMA-CTF(59) (POTENTIAL).
 FT CHAIN 695 751 GAMMA-CTF(57) (POTENTIAL).
 FT CHAIN 702 751 GAMMA-CTF(50) (POTENTIAL).
 FT CHAIN 721 751 C31 (POTENTIAL).
 FT DOMAIN 18 680 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 681 704 POTENTIAL.
 FT DOMAIN 705 751 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 96 110 HEPARIN-BINDING (BY SIMILARITY).
 FT DOMAIN 181 188 ZINC-BINDING (BY SIMILARITY).
 FT DOMAIN 291 341 BPTI/KUNITZ INHIBITOR.
 FT DOMAIN 316 344 HEPARIN-BINDING (BY SIMILARITY).
 FT DOMAIN 363 428 HEPARIN-BINDING (BY SIMILARITY).
 FT DOMAIN 504 521 COLLAGEN-BINDING (BY SIMILARITY).
 FT DOMAIN 713 732 INTERACTION WITH G(O)-ALPHA

FT				(BY SIMILARITY).
FT	DOMAIN	230	260	ASP/GLU-RICH (ACIDIC).
FT	DOMAIN	274	280	POLY-THR.
FT	SITE	144	144	REQUIRED FOR COPPER(II) REDUCTION
FT				(BY SIMILARITY).
FT	ACT_SITE	301	302	REACTIVE BOND.
FT	SITE	652	653	CLEAVAGE (BY BETA-SECRETASE)
FT				(BY SIMILARITY).
FT	SITE	653	654	CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT	SITE	668	669	CLEAVAGE (BY ALPHA-SECRETASE)
FT				(BY SIMILARITY).
FT	SITE	685	685	INVOLVED IN FREE RADICAL PROPAGATION
FT				(BY SIMILARITY).
FT	SITE	687	687	INVOLVED IN OXIDATIVE REACTIONS
FT				(BY SIMILARITY).
FT	SITE	692	693	CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT				(BY SIMILARITY).
FT	SITE	694	695	CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT				(BY SIMILARITY).
FT	SITE	701	702	CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)
FT				(BY SIMILARITY).
FT	SITE	705	715	BASOLATERAL SORTING SIGNAL
FT				(BY SIMILARITY).
FT	SITE	720	721	CLEAVAGE (BY CASPASES-3,-6,-8 OR -9)
FT				(BY SIMILARITY).
FT	SITE	738	741	ENDOCYTOSIS SIGNAL.
FT	SITE	740	743	NPXY MOTIF.
FT	METAL	137	137	COPPER (BY SIMILARITY).

Query Match 96.6%; Score 3921.5; DB 1; Length 751;
 Best Local Similarity 96.9%; Pred. No. 5.4e-192;
 Matches 746; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHMNVQNGKWDSDPSGTK	60
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHMNVQNGKWDSDPSGTK	60
Qy	61	TCIDTKEGILQYCQEVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG	120
Db	61	TCIDTKEGILQYCQEVPELQITNVVEANQPVTIQNWCKRDRKQCKTHPHIVIPYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE	240
Db	181	GVEFVCCPLAEESDHVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE	240
Qy	241	EADDDDEDDGDEVEEEAEPEYEEATERTTSIATTTTTTTESVEEVVREVCSEQAETGPC	300
Db	241	EADDDDEDDGDEVEEEAEPEYEEATERTTSIATTTTTTTESVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFVDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFVDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSV-----	344

QY 361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
 : |||||
 Db 345 ---IPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 401
 QY 421 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 |||||
 Db 402 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRLALENYITAL 461
 QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 |||||
 Db 462 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 521
 QY 541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGN DALMPSLTET 600
 |||||
 Db 522 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGN DALMPSLTET 581
 QY 601 KTTVELLPVNGEFSDDLQPWHSFGADSV PANTENEVEPV DARPAADRGLTTRPGSGLTN 660
 |||||
 Db 582 KTTVELLPVNGEFSDDLQPWHSFGADSV PANTENEVEPV DARPAADRGLTTRPGSGLTN 641
 QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 |||||
 Db 642 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 701
 QY 721 VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFE QMQN 770
 |||||
 Db 702 VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFE QMQN 751

RESULT 8

APP2_HUMAN

ID APP2_HUMAN STANDARD; PRT; 763 AA.
 AC Q06481;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Amyloid-like protein 2 precursor (Amyloid protein homolog) (APPH)
 DE (CDEI-box binding protein) (CDEBP).
 GN APLP2 OR APPL2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=93250009; PubMed=8485127;
 RA Sprecher C.A., Grant F.J., Grimm G., O'Hara P.J., Norris F.,
 RA Norris K., Foster D.C.;
 RT "Molecular cloning of the cDNA for a human amyloid precursor protein
 RT homolog: evidence for a multigene family.";
 RL Biochemistry 32:4481-4486(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=95217334; PubMed=7702756;
 RA von der Kammer H., Hanes J., Klaudiny J., Scheit K.H.;

RT "A human amyloid precursor-like protein is highly homologous to a
RT mouse sequence-specific DNA-binding protein.";
RL DNA Cell Biol. 13:1137-1143(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94035131; PubMed=8220435;
RA Wasco W., Gurubhagavatula S., Paradis M., Romano D.M., Sisodia S.S.,
RA Hyman B.T., Neve R.L., Tanzi R.E.;
RT "Isolation and characterization of APLP2 encoding a homologue of the
RT Alzheimer's associated amyloid beta protein precursor.";
RL Nat. Genet. 5:95-99(1993).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF HEMOSTASIS. THE
CC SOLUBLE FORM MAY HAVE INHIBITORY PROPERTIES TOWARDS COAGULATION
CC FACTORS. MAY INTERACT WITH CELLULAR G-PROTEIN SIGNALING PATHWAYS.
CC MAY BIND TO THE DNA 5'-GTCACATG-3' (CDEI BOX).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN AND NUCLEAR
CC (POTENTIAL).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=Q06481-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q06481-2; Sequence=VSP_000018;
CC Name=3;
CC IsoId=Q06481-3; Sequence=VSP_000019;
CC -!- TISSUE SPECIFICITY: IN PLACENTA, BRAIN, HEART, LUNG, LIVER, KIDNEY
CC AND ENDOTHELIAL TISSUES.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; S60099; AAC60589.1; -.
 DR EMBL; L09209; AAA35526.1; -.
 DR EMBL; Z22572; CAA80295.1; -.
 DR EMBL; L27631; AAC41701.1; -.
 DR EMBL; BC000373; AAH00373.1; -.
 DR PIR; A49321; A49321.
 DR HSSP; P05067; 1MWP.
 DR Genew; HGNC:598; APLP2.
 DR MIM; 104776; -.
 DR GO; GO:0016021; C:integral to membrane; NAS.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0003677; F:DNA binding activity; NAS.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; NAS.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.
 KW Transmembrane; Signal; Alternative splicing; DNA-binding;
 KW Nuclear protein; Serine protease inhibitor.
 FT SIGNAL 1 29 POTENTIAL.
 FT CHAIN 30 763 AMYLOID-LIKE PROTEIN 2.
 FT DOMAIN 30 692 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 693 716 POTENTIAL.
 FT DOMAIN 717 763 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 215 280 ASP/GLU-RICH (HIGHLY ACIDIC).
 FT DOMAIN 306 364 BPTI/KUNITZ INHIBITOR.
 FT DOMAIN 215 231 POLY-GLU.
 FT ACT_SITE 320 321 REACTIVE BOND (BY SIMILARITY).
 FT DISULFID 310 360 BY SIMILARITY.
 FT DISULFID 319 343 BY SIMILARITY.
 FT DISULFID 335 356 BY SIMILARITY.
 FT VARSPLIC 308 363 Missing (in isoform 2).
 FT /FTid=VSP_000018.
 FT VARSPLIC 613 624 Missing (in isoform 3).
 FT /FTid=VSP_000019.
 FT CONFLICT 543 543 S -> I (IN REF. 1).
 SQ SEQUENCE 763 AA; 86955 MW; CA3A7D6DDB8A28D0 CRC64;

Query Match 48.8%; Score 1981.5; DB 1; Length 763;
 Best Local Similarity 50.4%; Pred. No. 1.5e-93;
 Matches 408; Conservative 127; Mismatches 168; Indels 107; Gaps 21;

Qy 5 LALLLLAAWTARALEV-----PTDGNAG---LLAEPQIAMFCGRLNMHMNVQNGKWDSDP 56
 | | | | | | | : | | | : | | | | | | | : | | | : | | | : | |
 Db 15 LLLLLLVGLTAPALALAGYIEALANAGTGFVAEPQIAMFCGKLNMHVNIQTGKWEPPD 74
 Qy 57 SGTCTCIDTKEGILQYCQEVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYR 116
 : | | | : | | | : | | | | | : | | | | | : | | | : | | | : | | : | :
 Db 75 TGTKSCFETKEEVLQYCQEMYPELQITNVMEANQRVSIDNWCRRDKKQCKS--RFVTPFK 132
 Qy 117 CLVGEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGI 176
 | | | | | | | | | | : | | : | | : | | | | | | | : | : | | | | | :
 Db 133 CLVGEFVSDVLLVPEKCFHFKERMEVCENHQHWHTVVKACLTQGMTLYSYGMLLPCGV 192
 Qy 177 DKFRGVEFVCCPLAESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAE 236
 | : | | | : | | | : : | : | | : : | | | : | | | : | :
 Db 193 DQFHGT EYVCCPQTKIIGSVSKEEEEEDEE-----EEEEDEEEDYDVKSEFPTEAD 245
 Qy 237 VEE--EEA--DDDEDEDGDEVEEEAE-----EPYEEATERTTSIATTTTTTTTESVE 284
 : | : | | : | | | : | | | : | : : | | | : | | : | : : :
 Db 246 LEDFTEAAVDEDEDEDEEGEEVVEDRDYYDTFKGDDYNE--ENPTEPGSDGTMSDKEIT 303
 Qy 285 EVVREVCSEQAETGPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTTEYCMVAVCGSA 344
 | : | | : | | | | : | | | : | | | : | | | | | : | : | | | :
 Db 304 HDVKAVCSQEAMTGPCRAVMPRWYFDLSKGKCVRFIYGGCGGNRNNFESDYCMVAVCKAM 363
 Qy 345 MSQSLLKTTQEPLARDPVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRE 404
 : | | | | | | : | | | : | | | | | | : | | : | |
 Db 364 I-----PPTPLPT-----NDVDVVFETSADDNEHARFQKAKEQLEIRHRN 403
 Qy 405 RMSQVMREWEAAERQAKNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAM 464
 | | : | : | | | | | | | | : : : | | | | : | : | | | : | | | |
 Db 404 RMDRVKKEWEAAELQAKNLPKAERQTLIQHFQAMVKALEKEAASEKQQLVETHLARVEAM 463
 Qy 465 LNDRRRLALENYITALQAVPPRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAA 524
 | | | | : | | | : | | : | | | : | : | | | | | | : | : | | |
 Db 464 LNDRRRMALENYLAALQSDPPRPHRILQALRRYVRAENKDRLHTIRHYQHVLAVDPEKAA 523
 Qy 525 QIRSQVMTHLRVIYERMNQSLSLYNVPAVAEEIQDEVDLLOKEQNYSDDLANMISEP 584
 | : | | | | | | | | | | | | | | | | : | | | : | | | : | : | : |
 Db 524 QMKSQVMTHLHVIEERNQSLSLYKVPYVAQEIQEEIDELLQEQR-----ADM---- 572
 Qy 585 RISYGNDAIMPSTETKTTVELLPVNGEFSDDLQPWHSFGADSVPAANTENEVEPVDARP 644
 | | | | : | | | | | | : | | | : | : | | | | | | | | : | : |
 Db 573 -----DQFTASIS ETPVDVR---VSSEES-EEIPPFHPF--HPFPALPENE---DTQP 616
 Qy 645 AADRGLTTRPGSGLTN-----IKTEE---ISEVKMDAEFRHDSGYEVHHQKLVFFAEDVG 696
 : : | | : | | | : | : | : | : | : | : | : | : | | |
 Db 617 ELYHPM--KKGSGVGEQDGGGLIGAEKVINSKNKVDENMVIDETLDV--KEMIFNAERVG 672
 Qy 697 S-----NKGAIIGLMVGGVVIATVIVITLVMKKKQYTSIHHGVVEVD 739
 : | : | | : | | | | | : | | | : | | : | | | | |
 Db 673 GLEEEERESVGPLREDFSLSSSALIGLLVIAVAIATVIVISLVMKRKRQYGTISHGIVEVD 732
 Qy 740 AAVTPEERHLSKMQQNGYENPTYKFFEQQM 769
 : | | | | | : | | : | | | | : | | |
 Db 733 PMLTPEERHLNKMQNHYENPTYKYLEQQM 762

RESULT 9

APP2_RAT

ID APP2_RAT STANDARD; PRT; 765 AA.
AC P15943;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Amyloid-like protein 2 precursor (Sperm membrane protein YWK-II).
GN APLP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 1-627 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain, and Heart;
RX MEDLINE=94368849; PubMed=8086458;
RA Sandbrink R., Masters C.L., Beyreuther K.;
RT "Complete nucleotide and deduced amino acid sequence of rat amyloid
RT protein precursor-like protein 2 (APLP2/APPH): two amino acids length
RT difference to human and murine homologues.";
RL Biochim. Biophys. Acta 1219:167-170(1994).
RN [2]
RP SEQUENCE OF 575-765 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=90207205; PubMed=1690887;
RA Yan Y.C., Bai Y., Wang L.F., Miao S.Y., Koide S.S.;
RT "Characterization of cDNA encoding a human sperm membrane protein
RT related to A4 amyloid protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2405-2408(1990).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=A;
CC IsoId=P15943-1; Sequence=Displayed;
CC Name=B;
CC IsoId=P15943-2; Sequence=VSP_000021;
CC Name=C;
CC IsoId=P15943-3; Sequence=VSP_000020;
CC Name=D;
CC IsoId=P15943-4; Sequence=VSP_000020, VSP_000021;
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X77934; CAA54906.1; -.
DR EMBL; M31322; AAA42352.1; -.
DR PIR; A35981; A35981.
DR PIR; S42880; S42880.

```

DR      HSSP; P05067; 1MPW.
DR      InterPro; IPR001868; A4_APP.
DR      InterPro; IPR002223; Kunitz_BPTI.
DR      Pfam; PF02177; A4_EXTRA; 1.
DR      Pfam; PF00014; Kunitz_BPTI; 1.
DR      PRINTS; PR00203; AMYLOIDA4.
DR      PRINTS; PR00759; BASICPTASE.
DR      ProDom; PD000222; Kunitz_BPTI; 1.
DR      SMART; SM00006; A4_EXTRA; 1.
DR      SMART; SM00131; KU; 1.
DR      PROSITE; PS00319; A4_EXTRA; 1.
DR      PROSITE; PS00320; A4_INTRA; 1.
DR      PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR      PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW      Transmembrane; Alternative splicing; Serine protease inhibitor;
KW      Signal; Glycoprotein.
FT      SIGNAL      1      29      POTENTIAL.
FT      CHAIN      30      765      AMYLOID-LIKE PROTEIN 2.
FT      DOMAIN      30      695      EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM      696      718      POTENTIAL.
FT      DOMAIN      719      765      CYTOPLASMIC (POTENTIAL).
FT      DOMAIN      218      282      ASP/GLU-RICH (HIGHLY ACIDIC).
FT      DOMAIN      308      366      BPTI/KUNITZ INHIBITOR.
FT      ACT_SITE      322      323      REACTIVE BOND (BY SIMILARITY).
FT      DISULFID      312      362      BY SIMILARITY.
FT      DISULFID      321      345      BY SIMILARITY.
FT      DISULFID      337      358      BY SIMILARITY.
FT      DOMAIN      218      229      POLY-GLU.
FT      CARBOHYD      628      628      O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).
FT      VARSPLIC      311      365      Missing (in isoform C and isoform D).
FT      VARSPLIC      616      627      /FTId=VSP_000020.
FT      VARSPLIC      616      627      Missing (in isoform B and isoform D).
FT      VARSPLIC      616      627      /FTId=VSP_000021.
FT      CONFLICT      575      577      DQF -> EFV (IN REF. 2).
SQ      SEQUENCE      765 AA; 86882 MW; CF51FCCCE305A0CF CRC64;

```


RL Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE OF 1-246 FROM N.A.
 RX MEDLINE=94032480; PubMed=8218408;
 RA Hanes J., von der Kammer H., Kristjansson G.I., Scheit K.H.;
 RT "The complete cDNA coding sequence for the mouse CDEI binding
 RT protein.";
 RL Biochim. Biophys. Acta 1216:154-156(1993).
 RN [3]
 RP SEQUENCE OF 185-695 FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Heart;
 RX MEDLINE=93129193; PubMed=1482349;
 RA Vidal F., Blangy A., Rassoulzadegan M., Cuzin F.;
 RT "A murine sequence-specific DNA binding protein shows extensive local
 RT similarities to the amyloid precursor protein.";
 RL Biochem. Biophys. Res. Commun. 189:1336-1341(1992).
 RN [4]
 RP SEQUENCE OF 1-35 FROM N.A.
 RC STRAIN=129/Sv;
 RX MEDLINE=96029629; PubMed=7592716;
 RA von Koch C.S., Lahiri D.K., Mammen A.L., Copeland N.G.,
 RA Gilbert D.J., Jenkins N.A., Sisodia S.S.;
 RT "The mouse APLP2 gene. Chromosomal localization and promoter
 RT characterization.";
 RL J. Biol. Chem. 270:25475-25480(1995).
 CC -!- FUNCTION: BINDS TO THE DNA 5'-GTCACATG-3' (CDEI BOX) WHICH PLAYS
 CC AN IMPORTANT ROLE IN THE EARLY DEVELOPMENT OF EMBRYOS.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN AND NUCLEAR
 CC (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z22592; CAA80306.1; -.
 DR EMBL; M97216; AAA20039.1; -.
 DR EMBL; U34291; AAC52318.1; -.
 DR PIR; S38344; S38344.
 DR HSSP; P05067; 1MWP.
 DR MGD; MGI:88047; Aplp2.
 DR InterPro; IPR001868; A4_APP.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 KW Transmembrane; DNA-binding; Signal; Nuclear protein.
 FT SIGNAL 1 29 POTENTIAL.
 FT CHAIN 30 695 AMYLOID-LIKE PROTEIN 2.
 FT DOMAIN 30 624 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 625 648 POTENTIAL.
 FT DOMAIN 649 695 CYTOPLASMIC (POTENTIAL).

FT	DOMAIN	218	294	ASP/GLU-RICH (HIGHLY ACIDIC) .
FT	DOMAIN	218	231	POLY-GLU.
FT	DOMAIN	256	266	POLY-GLU.
FT	CARBOHYD	485	485	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CONFLICT	185	189	GMLLP -> MACCC (IN REF. 3) .
SQ	SEQUENCE	695 AA; 78944 MW; BBF4B95AAB2A0311 CRC64;		

Query Match 42.0%; Score 1704.5; DB 1; Length 695;
 Best Local Similarity 46.9%; Pred. No. 1.5e-79;
 Matches 371; Conservative 113; Mismatches 170; Indels 137; Gaps 24;

Qy	5	LALLLLAAWTARALEV-----PTDGNAG---LLAEPQIAMFCGRILNMHMNVQNGKWDS	56
		: : : : : :	
Db	15	LLVLLLLGLTAPAAALAGYIEALANAGTGFVAEAPQIAMLCGKLNMHVNIQTGWEPDP	74
Qy	57	SGTKTCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYR	116
		: : : : : : :	
Db	75	TGTKSCLGTKEEVLQYCQEIYPELQITNVMEANQPVNIDSWCRRDKRQCKS--HIVIPFK	132
Qy	117	CLVGFEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGI	176
		: : : : :	
Db	133	CLVGFEFVSDVLLVPDNCQFFHQERMEVCEKHQRWHTLVKEACLTGGLTLYSYGMLLPCGV	192
Qy	177	DKFRGVEFVCCPLAEESDNVDSADAEEEDSDVWGGADTDYADGSEDKVVEVAEEEEVAE	236
		: : : : : :	
Db	193	DQFHGTEYVCCP---QTKTVDS-----DSTMSK-----EEEE---	221
Qy	237	VEEEEADDDDEDEDGDEVEEEAEAPYEEATERTTSIATTTTTTTTESVEEVVREVCSEQAE	296
		: : : : : : :	
Db	222	-EEEEDEEEDYDLKSEFPTEADLEDFTE---AAADEEEEEDEEEGEEVVED-----	270
Qy	297	TGPCRAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEP	356
		: : : : : : :	
Db	271	-----RDYYD-----PF-----KGDDYNEENPTE-----PSSEGTI--SDKE	301
Qy	357	LARDPVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAA	416
		: : : : : :	
Db	302	IVHD-VKVPPTPLPTND-VDVYLETSADDNEHARFQKAKEQLEIRHRNRMDRVKKEWEEA	359
Qy	417	ERQAKNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRIALENY	476
		: : : : : :	
Db	360	ELQAKNLPKTERQTLIQHFQAMVKALEKEAASEKQQLVETHLARVEAMLNDRRRIALENY	419
Qy	477	ITALQAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRV	536
		: : : : : : : :	
Db	420	LAALQSDPPRPHRILQALRRYVRAENKDRLHTIRHYQHVLAVDPEKAAQMKSQVMTHLHV	479
Qy	537	IYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDAIMPS	596
		: : : : : : : :	
Db	480	IEERNQSLTLLYKVPYVAQEIQEEIDELLQEQR-----ADM-----DQFTSS	522
Qy	597	LTETKTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDPARPAADRGLTTRPGS	656
		: : : : : : : :	
Db	523	ISENPVDVRVSSEEESE-EIPPFHPLHPF-----PSLSENE-----GSGMAEQDG-	565
Qy	657	GLTNIKTEEI-SEVKMDAEFRHDSGYEVHHQKLVFFAEDVGS-----N	698
		: : : : : : : :	

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP POSSIBLE FUNCTION, AND TISSUE SPECIFICITY.
 RX MEDLINE=96115107; PubMed=7494461;
 RA Kim T.-W., Wu K., Xu J.-L., McAuliffe G., Tanzi R.E., Wasco W.,
 RA Black I.B.;
 RT "Selective localization of amyloid precursor-like protein 1 in the
 RT cerebral cortex postsynaptic density.";
 RL Brain Res. Mol. Brain Res. 32:36-44(1995).
 RN [5]
 RP HEPARIN AND ZINC BINDING.
 RX MEDLINE=95014513; PubMed=7929392;
 RA Bush A.I., Pettingell W.H. Jr., de Paradis M., Tanzi R.E., Wasco W.;
 RT "The amyloid beta-protein precursor and its mammalian homologues.
 RT Evidence for a zinc-modulated heparin-binding superfamily.";
 RL J. Biol. Chem. 269:26618-26621(1994).
 RN [6]
 RP INTERACTION WITH APBA2.
 RX MEDLINE=99107877; PubMed=9890987;
 RA Tomita S., Ozaki T., Taru H., Oguchi S., Takeda S., Yagi Y.,
 RA Sakiyama S., Kirino Y., Suzuki T.;
 RT "Interaction of a neuron-specific protein containing PDZ domains with
 RT Alzheimer's amyloid precursor protein.";
 RL J. Biol. Chem. 274:2243-2254(1999).
 RN [7]
 RP EXTRACELLULAR COPPER-BINDING.
 RX MEDLINE=22130992; PubMed=12135352;
 RA Simons A., Ruppert T., Schmidt C., Schlicksupp A., Pipkorn R.,
 RA Reed J., Masters C.L., White A.R., Cappai R., Beyreuther K.,
 RA Bayer T.A., Multhaup G.;
 RT "Evidence for a copper-binding superfamily of the amyloid precursor
 RT protein.";
 RL Biochemistry 41:9310-9320(2000).
 CC -!- FUNCTION: May play a role in postsynaptic function. The C-terminal
 CC gamma-secretase processed fragment, ALID1, activates transcription
 CC activation through APBB1 (Fe65) binding (By similarity). Couples
 CC to JIP signal transduction through C-terminal binding. May
 CC interact with cellular G-protein signaling pathways. Can regulate
 CC neurite outgrowth through binding to components of the
 CC extracellular matrix such as heparin and collagen I.
 CC -!- FUNCTION: The gamma-CTF peptide, C30, is a potent enhancer of
 CC neuronal apoptosis (By similarity).
 CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
 CC cytoplasmic proteins, including APBB and APBA family members,
 CC MAPK8IP1 and Dab1 (By similarity). Binding to Dab1 inhibits its
 CC serine phosphorylation (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. C-terminally
 CC processed in the Golgi complex.
 CC -!- TISSUE SPECIFICITY: Expressed in the cerebral cortex where it is
 CC localized to the postsynaptic density (PSD).

CC -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
 CC phosphorylated proteins is required for the specific binding of
 CC the PID domain. However additional amino acids either N- or C-
 CC terminal to the NPXY motif are often required for complete
 CC interaction. The NPXY site is also involved in clatherin-mediated
 CC endocytosis.
 CC -!- PTM: Proteolytically cleaved by caspases during neuronal
 CC apoptosis. Cleaved, in vitro, at Asp-620 by caspase-3 (By
 CC similarity).
 CC -!- PTM: N-glycosylated.
 CC -!- PTM: O-glycosylated.
 CC -!- MISCELLANEOUS: Binds zinc and copper in the extracellular domain.
 CC Zinc-binding increases heparin binding. No Cu(II) reducing
 CC activity with copper-binding.
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U48437; AAB96331.1; -.
 DR EMBL; AD000864; AAB50173.1; -.
 DR EMBL; BC012889; AAH12889.1; -.
 DR HSSP; P05067; 1MWP.
 DR Genew; HGNC:597; APLP1.
 DR MIM; 104775; -.
 DR GO; GO:0005604; C:basement membrane; TAS.
 DR GO; GO:0005208; F:amyloid protein; TAS.
 DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 DR GO; GO:0007399; P:neurogenesis; TAS.
 DR InterPro; IPR001868; A4_APP.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 KW Apoptosis; Endocytosis; Cell adhesion; Coated pits; Neurone;
 KW Heparin-binding; Metal-binding; Copper; Zinc; Signal; Transmembrane;
 KW Glycoprotein.
 FT SIGNAL 1 38 POTENTIAL.
 FT CHAIN 39 650 AMYLOID-LIKE PROTEIN 1.
 FT CHAIN 621 650 C30 (BY SIMILARITY).
 FT DOMAIN 39 580 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 581 603 POTENTIAL.
 FT DOMAIN 604 650 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 158 178 COPPER-BINDING (BY SIMILARITY).
 FT DOMAIN 204 211 ZINC-BINDING.
 FT DOMAIN 310 342 HEPARIN-BINDING (BY SIMILARITY).
 FT DOMAIN 410 441 HEPARIN-BINDING (BY SIMILARITY).
 FT DOMAIN 442 459 COLLAGEN-BINDING (BY SIMILARITY).
 FT DOMAIN 640 643 CLATHRIN-BINDING (POTENTIAL).
 FT DOMAIN 241 247 POLY-GLU.
 FT DOMAIN 264 268 POLY-GLU.

FT	SITE	167	167	REQUIRED FOR COPPER(II) REDUCTION (BY
FT				SIMILARITY).
FT	SITE	604	615	BASOLATERAL SORTING SIGNAL (BY
FT				SIMILARITY).
FT	SITE	620	621	CLEAVAGE (BY CASPASE-3) (BY SIMILARITY).
FT	SITE	638	641	ENDOCYTOSIS SIGNAL (BY SIMILARITY).
FT	SITE	640	643	NPXY MOTIF.
FT	CARBOHYD	337	337	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	461	461	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	551	551	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	48	48	A -> P (IN REF. 1).
SQ	SEQUENCE	650 AA; 72176 MW; B95F0F4D1C5CBAC7 CRC64;		

Query Match 28.6%; Score 1159.5; DB 1; Length 650;
 Best Local Similarity 35.7%; Pred. No. 6.2e-52;
 Matches 276; Conservative 114; Mismatches 229; Indels 155; Gaps 17;

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHMNVQNGKWDSDPSG	60
		: : : : : : : : :	
Db	23	LLPLLLLLLLRAQPAIGSLAGGSPGAAEAPGSAQVAGLCGRLLTLHRDLRTGRWE	82
Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHF-VIPYRCLV	119
		: : : : : : : : : : : : : : :	
Db	83	RCLRDPQRVLEYCRQMYPELQIARVEQATQAIPIMERWCGGSRSGSCAHPHHQVVPFRCLP	142
Qy	120	GEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKF	179
		: : : : : : :	
Db	143	GEFVSEALLVPEGCRFLHQERMDQCESSTRRHQEAQEAQSSQGLILHGSGMLLPCGSDRF	202
Qy	180	RGVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEE	239
		:	
Db	203	RGVEYVCCPPPGTPD--PSGTAVGDPSTRSW-----PPGSR---VEGAEDDEE---EE	246
Qy	240	EEADDDDEDDGDEVEEEAEPEYEEATERTTSIATTTTTTTESVEEVVREVCSEQAETGP	299
		: :	
Db	247	ESFPQPVDYDFVEPPQAEEDD-----ETVP	271
Qy	300	CRAMISRWFYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQEPLAR	359
		: :	
Db	272	-----PPSHTLAVVG-----KVT-----	285
Qy	360	DPVKLPTTAASTPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEAAERQ	419
Db	286	-PTPRPT-----DGVDIYFGMPGEISEHEGFLRAKMDLEERRMRQINEVMREWAMADNQ	338
Qy	420	AKNLPKADKKAVIQHFQEKVESLEQEAANERQQQLVETHMARVEAMLNDRRLALENYITA	479
		: : : : : : : : : : : : : : :	
Db	339	SKNLPKADRQALNEHFQSILQTLQEEQVSGERQRLVETHATRIVIALINDQRRAALEGFLAA	398
Qy	480	LQAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYE	539
		: : : : : : : : : : : : :	
Db	399	LQADPPQAERVLLALRRYLRAEQKEQRHTLRHYQHVAAVDPEKAQQMRQVHTHLQVIEE	458
Qy	540	RMNQSLSLLYNVPAAEEIQDEVDELLQKEQNYSDDLVLANMISEPRISYGNDAIMPSLTE	599
		: : : : : : : : : : : :	
Db	459	RVNQSLGLLDQNPHLAQELRPQIQELLHSEH-----LGPSELE	496

Qy 600 TKTTVELLPVNGEFLD--DLQPWHSFGADSVFANTENEVEPVDARPAADRGLTTRPGSG 657
 Db 497 A-----PAPGGSSEDKGGLQPPDS--KDDTFM-----TLPKGSTEQDAAS 534
 Qy 658 LTNIKTEEISEVKMDAEFRHDSGYEVHH---QKLVFFAEDVGSNKGAIIGLMVGGVVIAT 714
 Db 535 PEKEKMNPLEQYERKVNASVPRGFPHSSEIQRDELAPAGTGVVSREAVSGLLIMGAGGGS 594
 Qy 715 VIVITLVML-KKKQYTSIHGGVVEVDAAVTPEERHLSKMQONGYENPTYKFFEQ 767
 Db 595 LIVLSMLLLRRKKPYGAISHGVVEVDPMLTLEEQQRLRELQRHGYENPTYRFLLE 648

RESULT 12

APP1_MOUSE

ID APP1_MOUSE STANDARD; PRT; 653 AA.
 AC Q03157; Q8VC38;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Amyloid-like protein 1 precursor (APLP) (APLP-1) [Contains: C30].
 GN APLP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93066322; PubMed=1279693;
 RA Wasco W., Bupp K., Magendantz M., Gusella J.F., Tanzi R.E.,
 RA Solomon F.;
 RT "Identification of a mouse brain cDNA that encodes a protein related
 RT to the Alzheimer disease-associated amyloid beta protein precursor."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:10758-10762(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [3]

RP COLLAGEN-BINDING.

RX MEDLINE=96139497; PubMed=8576160;

RA Beher D., Hesse L., Masters C.L., Multhaup G.;

RT "Regulation of amyloid protein precursor (APP) binding to collagen and

RT mapping of the binding sites on APP and collagen type I.";

RL J. Biol. Chem. 271:1613-1620 (1996).

RN [4]

RP INTERACTION WITH DAB1.

RX MEDLINE=99389880; PubMed=10460257;

RA Homayouni R., Rice D.S., Sheldon M., Curran T.;

RT "Disabled-1 binds to the cytoplasmic domain of amyloid precursor-like

RT protein 1.";

RL J. Neurosci. 19:7507-7515 (1999).

RN [5]

RP INTERACTION WITH MAPK8IP1.

RX MEDLINE=21408156; PubMed=11517249;

RA Matsuda S., Yasukawa T., Homma Y., Ito Y., Niikura T., Hiraki T.,

RA Hirai S., Ohno S., Kita Y., Kawasumi M., Kouyama K., Yamamoto T.,

RA Kyriakis J.M., Nishimoto I.;

RT "C-jun N-terminal kinase (JNK)-interacting protein-1b/islet-brain-1

RT scaffolds Alzheimer's amyloid precursor protein with JNK.";

RL J. Neurosci. 21:6597-6607 (2001).

RN [6]

RP GAMMA-SECRETASE PROCESSING, INTERACTION WITH APBB1, AND MUTAGENESIS OF

RP TYR-641.

RX MEDLINE=22313598; PubMed=12228233;

RA Scheinfeld M.H., Ghersi E., Laky K., Fowlkes B.J., D'Adamio L.;

RT "Processing of beta-amyloid precursor-like protein-1 and -2 by gamma-

RT secretase regulates transcription.";

RL J. Biol. Chem. 277:44195-44201 (2002).

CC -!- FUNCTION: May play a role in postsynaptic function. The C-terminal

CC gamma-secretase processed fragment, ALID1, activates transcription

CC activation through APBB1 (Fe65) binding. Couples to JIP signal

CC transduction through C-terminal binding. May interact with

CC cellular G-protein signaling pathways. Can regulate neurite

CC outgrowth through binding to components of the extracellular

CC matrix such as heparin and collagen I.

CC -!- FUNCTION: The gamma-CTF peptide, C30, is a potent enhancer of

CC neuronal apoptosis (By similarity).

CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several

CC cytoplasmic proteins, including APBB and APBA family members,

CC MAPK8IP1 and Dab1 (By similarity). Binding to Dab1 inhibits its

CC serine phosphorylation.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein. C-terminally

CC processed in the Golgi complex.

CC -!- DOMAIN: The NPXY sequence motif found in many tyrosine-

CC phosphorylated proteins is required for the specific binding of

CC the PID domain. However additional amino acids either N- or C-

CC terminal to the NPXY motif are often required for complete

CC interaction. The NPXY site is also involved in clathrin-mediated

CC endocytosis.

CC -!- PTM: Proteolytically cleaved by caspases during neuronal

CC apoptosis. Cleaved, in vitro, at Asp-623 by caspase-3 (By

CC similarity).

CC -!- PTM: N-glycosylated.
CC -!- PTM: O-glycosylated.
CC -!- MISCELLANEOUS: Binds zinc and copper in the extracellular domain.
CC Zinc-binding increases heparin binding. No Cu(II) reducing
CC activity with copper-binding.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L04538; AAA37247.1; -.
DR EMBL; BC021877; AAH21877.1; -.
DR PIR; A46362; A46362.
DR HSSP; P05067; 1MWP.
DR MGD; MGI:88046; Aplp1.
DR InterPro; IPR001868; A4_APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00006; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
KW Apoptosis; Endocytosis; Cell adhesion; Coated pits; Neurone;
KW Heparin-binding; Metal-binding; Copper; Zinc; Signal; Transmembrane;
KW Glycoprotein.
FT SIGNAL 1 37 POTENTIAL.
FT CHAIN 38 653 AMYLOID-LIKE PROTEIN 1.
FT CHAIN 624 653 C30 (BY SIMILARITY).
FT DOMAIN 38 583 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 584 606 POTENTIAL.
FT DOMAIN 607 653 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 157 177 COPPER-BINDING.
FT DOMAIN 203 210 ZINC-BINDING (BY SIMILARITY).
FT DOMAIN 313 345 HEPARIN-BINDING (BY SIMILARITY).
FT DOMAIN 413 444 HEPARIN-BINDING (BY SIMILARITY).
FT DOMAIN 445 462 COLLAGEN-BINDING (BY SIMILARITY).
FT DOMAIN 263 271 POLY-GLU.
FT DOMAIN 535 538 POLY-SER.
FT DOMAIN 601 606 POLY-LEU.
FT SITE 166 166 REQUIRED FOR COPPER(II) REDUCTION (BY
FT SIMILARITY).
FT SITE 607 618 BASOLATERAL SORTING SIGNAL (BY
FT SIMILARITY).
FT SITE 623 624 CLEAVAGE (BY CASPASE-3) (BY SIMILARITY).
FT SITE 641 644 ENDOCYTOSIS SIGNAL.
FT SITE 643 646 NPXY MOTIF.
FT CARBOHYD 464 464 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 554 554 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MUTAGEN 641 641 Y->G: REDUCED BINDING OF APBB1.
FT CONFLICT 17 17 P -> PP (IN REF. 2).
SQ SEQUENCE 653 AA; 72750 MW; 56516DC3EA40E4B0 CRC64;

Query Match

28.4%; Score 1150.5; DB 1; Length 653;

Best Local Similarity 35.4%; Pred. No. 1.8e-51;
Matches 274; Conservative 120; Mismatches 228; Indels 153; Gaps 19;

Qy	1	MLPGLALLLLAAWTARA-LEVPTDGNAGLLAEPQIAMFCGRGLNMHMNVQNGKWDSDPSGT	59
Db	22	LLP-LSLLLLRAQLAVGNLAVGSPSAAEAPGSAQVAGLCGRLLTLHRDLRTGRWEPDPQRS	80
Qy	60	KTCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHF-VIPYRCL	118
Db	81	RRCLLDPQRVLEYCRQMPYELHIARVEQAAQAI PMERWCGGTRSGRCAHPHHEVVPFHCL	140
Qy	119	VGEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDK	178
Db	141	PGEFVSEALLVPEGCRFLHQERMDQCESSTRRHQEAQEACSSQGLILHSGSMLLPCGSDR	200
Qy	179	FRGVEFVCCPLAEESDNVDSADAEEDDSDVW-WGGADTDYADGSEDKVVEVAEEEEVAEV	237
Db	201	FRGVEYVCCP-PPATPNPSGMAAGDPSTRSWPLGGR----AEGGED-----EEEVESF	248
Qy	238	EEEEADDDDEDDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVVEEVVREVCSEQAET	297
Db	249	PQPVDYDFVEPPQAEEEEEEEEEERAPPPSSHTP-----	281
Qy	298	GPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNNDTEEYCMAVCGSAMSQSLLKTTQEPL	357
Db	282	----VMVSR----VT-----	288
Qy	358	ARDPVKLPTTAASTPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEAE	417
Db	289	---PTPRPT-----DGVDVYFGMPGEIGEHEGFLRAKMDLEERRMRQINEVMREWAMAD	339
Qy	418	RQAKNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYI	477
Db	340	SQSKNLPKADRQALNEHFQSILQTL EEQVSGERQRLVETHATRVIALLINDQRRAALEGFL	399
Qy	478	TALQAVPPRPRHVFNMMLKKYVRAEQKDRQHTLKHFEHVRMVPDKKAAQIRSQVMTHLRVI	537
Db	400	AALQGDPQAEERVLMA LRRYLRAEQKEQRHTLRHYQHVAVDPEKAQQMRQVQTHLQVI	459
Qy	538	YERMNQSLSLLYNVPAAVEEIQDEVDLLQKEQNYSDDLANMISEPRISYGN DALMP-S	596
Db	460	EERMNQSLGLLDQNPHLAQELRPQIQELL-----LAEHLGPSEL----DASVPGS	505
Qy	597	LTETKTTVELLPVNGEFLSDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGS	656
Db	506	SSEDK-----GSLQP-----PESKDDPPVTLP---KGSTDQESS	536
Qy	657	GLTNIKTEEISEVKMDAEFRHDSGYEVHH---QKLVFFAEDVGSNKGAIIGLMVGGVVIA	713
Db	537	SSGREKLTPLEQYEQVNASAPRGFPFHSSDIQRDELAPSGTGVSREALSGLLIMGAGGG	596
Qy	714	TVIVITLVM L-KKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQ	767
Db	597	SLIVLSLLLLRKKKPYGTISHGVVEVDPM LTL EEQQLRELQRHGYENPTYRFL EE	651

A4_CAEEL
ID A4_CAEEL STANDARD; PRT; 686 AA.
AC Q10651; Q18583; Q95ZX1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Beta-amyloid-like protein precursor.
GN APL-1 OR C42D8.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE OF 6-686 FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94089766; PubMed=8265668;
RA Daigle I., Li C.;
RT "apl-1, a Caenorhabditis elegans gene encoding a protein related to
RT the human beta-amyloid protein precursor.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:12045-12049(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Hallsworth K.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Waterston R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=a;
CC IsoId=Q10651-1; Sequence=Displayed;
CC Name=b;
CC IsoId=Q10651-2; Sequence=VSP_000017;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00240; AAC46470.1; ALT_INIT.
DR EMBL; U56966; AAA98722.1; -.
DR EMBL; U56966; AAK68242.1; -.
DR PIR; T15795; T15795.
DR HSSP; P05067; 1MWP.
DR WormPep; C42D8.8a; CE04209.
DR WormPep; C42D8.8b; CE27845.
DR InterPro; IPR001868; A4_APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR PRINTS; PR00203; AMYLOIDA4.

DR SMART; SM00006; A4_EXTRA; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 KW Signal; Transmembrane; Amyloid; Neurogenesis; Glycoprotein;
 KW Alternative splicing.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 686 BETA-AMYLOID-LIKE PROTEIN.
 FT DOMAIN 22 621 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 622 642 POTENTIAL.
 FT DOMAIN 643 686 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 205 228 ASP-RICH.
 FT DOMAIN 676 679 CLATHRIN-BINDING (POTENTIAL).
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 249 249 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 538 539 Missing (in isoform b).
 FT /FTid=VSP_000017.
 SQ SEQUENCE 686 AA; 79434 MW; A0816858FDD48608 CRC64;

Query Match 19.3%; Score 785; DB 1; Length 686;
 Best Local Similarity 26.8%; Pred. No. 6.7e-33;
 Matches 224; Conservative 111; Mismatches 272; Indels 230; Gaps 23;

QY 1 MLPG LALLLLA AWTARALEVPTDGNAGLLAEPQIAMFCGR LNMH MNVQNGKWDSDPSG TK 60
 :: || : : | | | | | : | | | : | : | : |
 Db 6 LMIGLLIPILVA-TVYAEGSPAGSKRHEKFI PMVAFSCGYRNQYM-TEEGSWKTDDERYA 63
 QY 61 TCIDTKEGILQYCQEVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
 || | ||: ||: || : |||: || : |: | : ||: || | | || | : |
 Db 64 TCFSGKLDILKYCRKAYPSMNITNIVEYSHEVSI SDWCREEGSPCK-WTHSVRPYHCIDG 122
 QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTN-----LHDYGM LLLPC 174
 || |: || | | |: | | | | : || | : | | : : : | ||
 Db 123 EFHSEALQVPHDCQF SHVNSRDQCNDYQHWKDEAGKQCKTKKSKGNKDMI VRSFVLEPC 182
 QY 175 GIDKFRGVEFVCCPLAEE SDNVDSADA EEDSDSVWGGADTDYADGSEDKVVEVAEEEEV 234
 : | | ||||| | : | : | : :
 Db 183 ALDMFTGVEFVCCP----NDQTNKTDVQKTK----- 209
 QY 235 AEVEEEEADDDDEDEDGDEVEEEAE EEPYEEATERTTSIATTTTTTTTESVEEVVREVCSEQ 294
 |: |: ||||| |: ||: | ||
 Db 210 ---EDEDDEDDDAYEDDYSEESDEKDEE----- 236
 QY 295 AETGPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCM AVCGSAMSQSLLKTTQ 354
 Db 237 ----- 236
 QY 355 EPLARDPVKLPTTAASTPDAVDKYLET PGDENEHAHFQKAKERLEAKHRERMSQVMREWE 414
 || : : | | : | | | |: |: |: | | : : | : |
 Db 237 EPSSQDP-----YFKIANWTNEHDDFKKAEMRMDEKHKRKKVDKVMKEWG 280
 QY 415 EA-----ERQAKNLPKADKKAVIQ---HFQEKVESLEQE AANERQQLVETHMARVEAMLN 466
 : |: |: | | : | | | |: | | |: | : : | | : | |
 Db 281 DLETRYNEQKAKD-PKGAEKFKSQMNARFQKT VSSLEEEHKRMRKEI EAVHEERVQAMLN 339
 QY 467 DRRRLALENYITAL--QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAA 524
 : : | | : | | | | | | | |: |: | | | : | : | | : | |

Db 340 EKKRDATHDYRQALATHVKNPKHNSVLQSLKAYIRAEKDRMHTLNRYRHLLKADSKEAA 399

Qy 525 QIRSQVMTHLRVIYERMNQSLSLLYNVP-----AVA--EEIQDEVDELLQKEQNYS 573
: | : | | | : | : | : | : | : | : | : | : | : |

Db 400 AYKPTVIHRLRYIDLRLINGTLAMLRDFPDLEKYVRPIAVTYWKDYRDEVSPDISVE---- 455

Qy 574 DDVLANMISEPRISYGN--DALMPSLT----ETKTTVELLPVNGEFSLDDLQPWHSFGAD 627
| | : | : | : | : | : | : | : | : | : | : |

Db 456 DSELTPIIHDEFKSKNAKLDVKAPTTTAKPVKETDNKVLPTASDSEEEADEYYEDED 515

Qy 628 SVPANT---ENEVEPVDPARP-----AADRGLTTRPGSGLTNIKT 663
| : : | : | : | : | : | : | : | : | : |

Db 516 EQVKKTPDMKKKVKVVDIKPKEIKVTIEEEKKAPKLVETSVQTDDEDDDDSSSSSTSSSES 575

Qy 664 EE-----ISEVKMDAE-----FRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGG 709
: | : | : | : | : | : | : | : | : | : | : |

Db 576 DEDEDKNIKELRVDIEPIIDEPASFYRHD-----KLIQSPEVERSASSVFQPYVLAS 627

Qy 710 VVIATVIVITLVMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFE 766
: | | | : : : | : | : | : | : | : | : | : | : |

Db 628 AMFITAICIIAFAITNARRRRAMRGFIEVD-VYTPEERHVAGMQVNGYENPTYSFDD 683

RESULT 14

A4_DROME

ID A4_DROME STANDARD; PRT; 887 AA.

AC P14599; Q9TVV0; Q9U4H3; Q9W5F1;

DT 01-APR-1990 (Rel. 14, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Beta-amyloid-like protein precursor.

GN APPL OR VND OR BCDNA:GH04413 OR EG:65F1.5 OR CG7727.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89184650; PubMed=2494667;

RA Rosen D.R., Martin-Morris L., Luo L., White K.;

RT "A Drosophila gene encoding a protein resembling the human

RT beta-amyloid protein precursor.";

RL Proc. Natl. Acad. Sci. U.S.A. 86:2478-2482(1989).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP REVISIONS.
 RC STRAIN=Berkeley;
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=20196011; PubMed=10731137;
 RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
 RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demailles J., Cadieu E.,
 RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,
 RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
 RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
 RA Modolell J., Peter A., Schoettler P., Werner M., Mourkioti F.,
 RA Beinert N., Dowe G., Schaefer U., Jaekle H., Bucheton A.,
 RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,
 RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
 RA Glover D.M.;

RT "From sequence to chromosome: the tip of the X chromosome of D.
 RT melanogaster.";
 RL Science 287:2220-2222(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley; TISSUE=Ovary;
 RX MEDLINE=20196012; PubMed=10731138;
 RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E.,
 RA Stapleton M., Harvey D.A.;
 RT "A Drosophila complementary DNA resource.";
 RL Science 287:2222-2224(2000).
 RN [6]
 RP SEQUENCE OF 1-83 FROM N.A.
 RX MEDLINE=91184006; PubMed=2127912;
 RA Martin-Morris L.E., White K.;
 RT "The Drosophila transcript encoded by the beta-amyloid protein
 RT precursor-like gene is restricted to the nervous system.";
 RL Development 110:185-195(1990).
 CC -!- FUNCTION: Probably corresponds to the protein encoded by the
 CC essential locus vnd, a gene required for embryonic nervous system
 CC development.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed in post-mitotic neurons in the
 CC central and peripheral nervous systems. Within the nervous system
 CC transcripts are not observed in neuroblasts, newly generated
 CC neurons and at least one class of presumed glial cells.
 CC -!- DEVELOPMENTAL STAGE: Expressed in all developmental stages.
 CC -!- DOMAIN: The clathrin-binding site is essential for its association
 CC with x11-alpha, -beta, and -gamma. The sequence specific
 CC recognition extends to peptide residues that are C-terminal to the
 CC NPXY motif. This interaction appears to be independent of
 CC phosphorylation (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J04516; AAA28874.1; -.
 DR EMBL; AE003418; AAF45520.2; -.
 DR EMBL; AL031883; CAA21409.1; -.
 DR EMBL; AL022139; CAA21409.1; JOINED.
 DR EMBL; AL022139; CAA18093.1; -.
 DR EMBL; AL031883; CAA18093.1; JOINED.
 DR EMBL; AF181628; AAD55414.1; -.
 DR EMBL; X55774; CAA39294.1; -.
 DR EMBL; X55775; CAA39294.1; JOINED.
 DR PIR; A32758; A32758.
 DR HSSP; P05067; 1MWP.
 DR FlyBase; FBgn0000108; Appl.
 DR GO; GO:0005576; C:extracellular; IDA.
 DR GO; GO:0005886; C:plasma membrane; IDA.
 DR InterPro; IPR001868; A4_APP.

DR Pfam; PF02177; A4_EXTRA; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 KW Signal; Transmembrane; Amyloid; Neurogenesis.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 887 BETA-AMYLOID-LIKE PROTEIN.
 FT DOMAIN 28 813 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 814 834 POTENTIAL.
 FT DOMAIN 835 887 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 877 880 CLATHRIN-BINDING (POTENTIAL).
 FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 574 574 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 177 177 S -> T (IN REF. 1).
 FT CONFLICT 229 229 MISSING (IN REF. 1).
 FT CONFLICT 332 332 V -> M (IN REF. 4).
 FT CONFLICT 743 743 S -> T (IN REF. 1).
 SQ SEQUENCE 887 AA; 98332 MW; F0F0855AD65A5275 CRC64;

Query Match 18.2%; Score 739; DB 1; Length 887;
 Best Local Similarity 24.7%; Pred. No. 2e-30;
 Matches 235; Conservative 138; Mismatches 312; Indels 266; Gaps 30;

Qy 7 LLLLLAAWTARALEVPTDGNAGLLA-----EPQIAMFC--GRLNMHMNV-QNGKWDSDPSG 58
 ||| : | | : | | : | |||| : | | : : : | : | |
 Db 9 LLLRSLWVVLAI-----GTAQVQAASPRWEPQIAVLCEAGQIQPQYLSEEGRWVTDLSK 63
 Qy 59 T---KTCIDTKEGILQYCQEVPELQITNVVEANQPVTIQNWCKRG---RKQCKTHPHFV 112
 || : | : | || : || ||| : || : | || : | : || : :
 Db 64 KTTGPTCLRDKMDLLDYCKKAYPNRDI TNIVESSHYQKIGGWCRQGALNAAKCKGSHRWI 123
 Qy 113 IPYRCLVGEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMILL 172
 | : ||| | | ||||| : | | | | : | : | : : | : |||
 Db 124 KPFRCCL-GPFQSDALLVPEGCLFDHIHNASRCWPFVRWNQTGAAACQERGMQMRSFAMLL 182
 Qy 173 PCGIDKFRGVEFVCCP-----LAESDNVD---S 198
 |||| | ||||| ||| : : | : | |
 Db 183 PCGISVFSGVFVCCPKHFKTDEIHVKKTDLPVMPAAQINSANDELVMNDEDDSDNSNYS 242
 Qy 199 ADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEV-----AE 236
 || ||| | | | : : || : | ||
 Db 243 KDANEDDLLD-----DEDDLMDGDEEDDMVADEAATAGGSPNTGSSGDSNSGSLDDINAE 296
 Qy 237 VEE-EEADDDDEDEDGDEVEEEAEPEYEEA-TERTTSIATTTTTTTESVEEVVREVCSEQ 294
 : || | : | | | | : : : : | : : : : : : :
 Db 297 YDSGEEGDNYEEDGAGSESEAEVEASWDQSGGAKVVSLSKSDSSSPSSAPVAPAPEKAPVK 356
 Qy 295 AETGPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQ 354
 : | : : | | | | | | : | : | : |
 Db 357 SESVTSTPQLS-----ASAAAFVAANSNGSGT-----GAGAPPSTAQPTS 396
 Qy 355 EPLARDPVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWE 414
 : | | | | | | : : : : ||| ||| : : : :
 Db 397 DP-----YFTHFDPHYEHQSYKVSQKRLEESHREKVTRVMKDW 435

Qy	415	EAERQAKNLPKADKKA-----VIQHFQEKVESLEQEAANRQQQLVETHMARVEAMLND	467
Db	436	DLEEKYQDMRLADPKAAQSFQQRMTARFQTSVQALEEEGNAEKHQLAAMHQQRVLAHINQ	495
Qy	468	RRRLALENYITALQAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEH-VRMVDP---KKA	523
Db	496	RKREAMTCYTQALTEQPPNAHHVEKCLQKLLRALHKDRAHALAHYRHLNLSGGPGGLEAA	555
Qy	524	AQIRSQVMTHLRVIYERMNQSLSLLYNVPAAVEEI-----QDEV-----	562
Db	556	ASERPRTLRLIDIDRAVNQSMTMLKRYPELSAKIAQLMNDYIILALRSKDDIPGSSLGMS	615
Qy	563	-----DELLQKEQNYSDDVLA	578
Db	616	EEAEAGILDKYRVEIERKVAEKERLRLAEKQQRKEQRAAEREKLRREEKLRLEAKKVDDMLK	675
Qy	579	NMISE-----PRISYGNDALM-----PSLTETKTTVELLPVN	610
Db	676	SQVAEQQSQPTQSSTQSQAQQQQQEKSLPGKELGPDAAALVTAANPNLETTKS-----	727
Qy	611	GEFSLDDLQPWHSFGADSVPAANTENEVEPVDPARPAADRGLTTRPGSGLTNIKTEEISEVK	670
Db	728	-EKDLSDETE---YGEATVSSTKVQTVLPTVDDDAVQRAVEDVAAA-----VAHQE	773
Qy	671	MDAEFRHDSGYEVHHQKLVF-----FAEDVGSNK---GAIIGLMVGGVVIATVIVITLVM	722
Db	774	AEPQVQHFMTHDLGHRESSFSLRREFAQHAHAAKEGRNVYFTLSFAGIALMAAVFVGAV	833
Qy	723	LKKKQYTSIH-HGVVEVDAAVTP-----EERHLSKMQQNGYENPTYKFFE	766
Db	834	AKWRTSRSPHAQGFIEVDQNVTTTHHPIVREEKIVPNMQINGYENPTYKYFE	884

A4 MACMU

AC P29216;

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DE Alzheimer's disease amyloid A4 protein (Fragm

GN APP.

OS Maca

OC Eukaryota; Metazoa; Chordata; Cr

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;

OC Cercopithecinae; Macaca.

OX NCBI TaxID=9544;

RN [1] —

RP SEQ

RC TISSUE=Brain;

RX MEDLINE=90180

RA Koo E.H., Sisodia S.S., Cork L.C.

RA Price D.L.;

RT "Differenti

RT of Alzheimer's disease and in aged nonhuman primates.";

RL Neuron 4:97-104 (1990).

CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
 CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
 CC G(O) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=5;
 CC Comment=Experimental confirmation may be lacking for some
 CC isoforms;
 CC Name=APP(770);
 CC IsoId=P29216-1; Sequence=Displayed;
 CC Name=APP(395);
 CC IsoId=P29216-2; Sequence=Not described;
 CC Name=APP(563);
 CC IsoId=P29216-3; Sequence=Not described;
 CC Name=APP(695);
 CC IsoId=P29216-4; Sequence=Not described;
 CC Name=APP(751);
 CC IsoId=P29216-5; Sequence=Not described;
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X15985; CAA34116.1; -.
 DR PIR; S06678; S06678.
 DR HSSP; P05067; 1AAP.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; PARTIAL.
 DR PROSITE; PS00320; A4_INTRA; PARTIAL.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.
 KW Glycoprotein; Amyloid; Neurone; Alternative splicing;
 KW Serine protease inhibitor.
 FT NON_TER 1 1
 FT DOMAIN 1 76 BPTI/KUNITZ INHIBITOR.
 FT ACT_SITE 13 14 REACTIVE BOND.
 FT DISULFID 3 53 BY SIMILARITY.
 FT DISULFID 12 36 BY SIMILARITY.
 FT DISULFID 28 49 BY SIMILARITY.
 FT NON_TER 76 76
 SQ SEQUENCE 76 AA; 8527 MW; 492BF3069AB082A1 CRC64;

Query Match 9.9%; Score 403; DB 1; Length 76;
 Best Local Similarity 94.7%; Pred. No. 1.1e-14;
 Matches 72; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

Qy      289 EVCSEQAETGPCRAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQS 348
        |||
Db      1   EVCSEQAETGPCRAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSVMSQS 60
        |||

Qy      349 LLKTTQEPLARDPVKL 364
        |||:|||||
Db      61 LRKTTREPLTRDPVKL 76

```

Search completed: December 13, 2003, 13:35:04
Job time : 30.3248 secs

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 47.9487 Seconds
(without alignments)
1544.356 Million cell updates/sec

Title: US-09-785-215-2
Perfect score: 4058
Sequence: 1 MLPGLALLLLAAWTARALEV.....KMQQNGYENPTYKFFEQMQN 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query Match	Length			
1	4058	100.0	770	1	QRHUA4	Alzheimer's diseas
2	3590.5	88.5	695	1	A49795	Alzheimer's diseas
3	3493.5	86.1	695	2	S00550	Alzheimer's diseas
4	3468.5	85.5	695	2	A27485	Alzheimer's diseas
5	3403.5	83.9	747	2	JH0773	Alzheimer's diseas
6	2446.5	60.3	484	4	A32761	hypothetical Alzhe
7	1981.5	48.8	763	2	A49321	amyloid beta (A4)
8	1971.5	48.6	765	2	S42880	amyloid precursor-
9	1956.5	48.2	751	2	A49974	beta-amyloid precu
10	1150.5	28.4	653	2	A46362	amyloid precursor-
11	1112.5	27.4	511	2	JC1404	CDEI-box DNA-bindi
12	785	19.3	686	2	T15795	hypothetical prote
13	737.5	18.2	886	2	A32758	beta-amyloid-like

14	706	17.4	246	2	S38344	CDEI-binding prote
15	501	12.3	100	2	A32282	Alzheimer's diseas
16	411	10.1	82	2	PQ0438	Alzheimer's diseas
17	403	9.9	76	2	S06678	Alzheimer's diseas
18	397	9.8	76	2	S03607	Alzheimer's diseas
19	383	9.4	76	2	S04855	Alzheimer's diseas
20	296.5	7.3	191	2	A35981	sperm membrane pro
21	283	7.0	57	2	E60045	Alzheimer's diseas
22	283	7.0	57	2	F60045	Alzheimer's diseas
23	283	7.0	57	2	G60045	Alzheimer's diseas
24	283	7.0	57	2	D60045	Alzheimer's diseas
25	283	7.0	57	2	A60045	Alzheimer's diseas
26	283	7.0	57	2	B60045	Alzheimer's diseas
27	256.5	6.3	111	2	S41082	amyloid precursor
28	217	5.3	42	2	PN0512	beta-amyloid prote
29	194.5	4.8	2225	2	T26063	hypothetical prote
30	186	4.6	993	2	S49461	synaptonemal compl
31	183.5	4.5	396	2	S53325	tissue factor path
32	183	4.5	1188	2	T46608	zinc finger protei
33	179.5	4.4	1208	2	T27822	hypothetical prote
34	178.5	4.4	1110	2	I51116	NF-180 - sea lampr
35	177.5	4.4	2167	2	T34395	hypothetical prote
36	176.5	4.3	252	2	JG0185	hepatocyte growth
37	176	4.3	55	2	S30332	proteinase inhibit
38	175.5	4.3	579	2	JH0820	160K golgi antigen
39	175	4.3	302	1	TIRTGK	tissue factor path
40	175	4.3	1182	2	T30189	myelin transcripti
41	174.5	4.3	1271	2	A45555	glutamate rich pro
42	173.5	4.3	1087	2	T30330	gelsolin-related p
43	173.5	4.3	1558	2	C89114	protein C37C3.6a [
44	172	4.2	922	2	T23573	hypothetical prote
45	170.5	4.2	5170	2	T15348	hypothetical prote

ALIGNMENTS

RESULT 1

QRHUA4

Alzheimer's disease amyloid beta protein precursor [validated] - human

N;Alternate names: Alzheimer's disease amyloid A4 protein; coagulation factor XIa inhibitor; proteinase nexin II (PN-II)

N;Contains: amyloid beta protein long, plaque form; amyloid beta protein short, vascular form; amyloid protein precursor splice form APP(695); amyloid protein precursor splice form APP(751); amyloid protein precursor splice form APP(770)

C;Species: Homo sapiens (man)

C;Date: 30-Jun-1987 #sequence_revision 28-Jul-1995 #text_change 15-Sep-2000

C;Accession: S02260; S05194; A32277; A33260; A35486; I39452; I39451; I39453; I59562; A44017; B44017; A03134; A29030; A47584; A47585; S02638; S00707; S00925; A38949; A30320; B30320; C30320; A31087; A24668; A28583; A29302; A60805; JL0038; S06121; A60355; A59011; A38384; S29076; S38252; S32539; S48148; S48692; S51186; S51185; S51184; S51183; A54238; I58075; I52250; S09010; S10737; S24127; S43644

R;Lemaire, H.G.; Salbaum, J.M.; Multhaup, G.; Kang, J.; Bayney, R.M.; Unterbeck, A.; Beyreuther, K.; Mueller-Hill, B.

Nucleic Acids Res. 17, 517-522, 1989

A;Title: The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid is encoded by 16 exons.

A;Reference number: S02260; MUID:89128427; PMID:2783775
 A;Accession: S02260
 A;Molecule type: DNA
 A;Residues: 1-288,'V',365-770 <LEM1>
 A;Cross-references: EMBL:X13466
 A;Note: alternative splice form APP(695)
 R;Lemaire, H.G.
 submitted to the EMBL Data Library, November 1988
 A;Reference number: S05194
 A;Accession: S05194
 A;Molecule type: DNA
 A;Residues: 1-14,'VW',17-288,'V',365-770 <LEM2>
 A;Cross-references: EMBL:X13466; NID:g35598; PIDN:CAA31830.1; PID:g871360
 A;Note: alternative splice form APP(695)
 R;La Fauci, G.; Lahiri, D.K.; Salton, S.R.J.; Robakis, N.K.
 Biochem. Biophys. Res. Commun. 159, 297-304, 1989
 A;Title: Characterization of the 5'-end region and the first two exons of the
 beta-protein precursor gene.
 A;Reference number: A32277; MUID:89165870; PMID:2538123
 A;Accession: A32277
 A;Molecule type: DNA
 A;Residues: 1-75 <LAF>
 A;Cross-references: GB:M24546; GB:M24547; NID:g341202; PIDN:AAC13654.1;
 PID:g516074
 R;Johnstone, E.M.; Chaney, M.O.; Moore, R.E.; Ward, K.E.; Norris, F.H.; Little,
 S.P.
 Biochem. Biophys. Res. Commun. 163, 1248-1255, 1989
 A;Title: Alzheimer's disease amyloid peptide is encoded by two exons and shows
 similarity to soybean trypsin inhibitor.
 A;Reference number: A33260; MUID:89392030; PMID:2675837
 A;Accession: A33260
 A;Molecule type: DNA
 A;Residues: 656-737 <JOH>
 A;Cross-references: GB:M29270; NID:g178863; PIDN:AAA51768.1; PID:g178865
 R;Prelli, F.; Levy, E.; van Duinen, S.G.; Bots, G.T.A.M.; Luyendijk, W.;
 Frangione, B.
 Biochem. Biophys. Res. Commun. 170, 301-307, 1990
 A;Title: Expression of a normal and variant Alzheimer's beta-protein gene in
 amyloid of hereditary cerebral hemorrhage, Dutch type: DNA and protein
 diagnostic assays.
 A;Reference number: A35486; MUID:90321244; PMID:2196878
 A;Accession: A35486
 A;Molecule type: DNA
 A;Residues: 672-710 <PRE1>
 A;Note: 693-Gln was found in DNA isolated from HCHWA-D patients
 R;Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
 Gene 87, 257-263, 1990
 A;Title: Genomic organization of the human amyloid beta-protein precursor gene.
 A;Reference number: I39451; MUID:90236318; PMID:2110105
 A;Accession: I39452
 A;Status: nucleic acid sequence not shown; translation not shown; translated
 from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-770 <YOS1>
 A;Cross-references: GB:M33112; NID:g178613; PIDN:AAB59502.1; PID:g178616
 A;Accession: I39451

A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-530, 'QWLMPVIPAFWEAKVGR' <YOS2>
 A;Cross-references: GB:M34875; NID:g178608; PIDN:AAB59501.1; PID:g178615
 R;Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
 Gene 102, 291-292, 1991
 A;Reference number: A59020; MUID:91340168; PMID:1908403
 A;Contents: annotation; erratum
 A;Note: revised physical map for reference I39451
 R;Levy, E.; Carman, M.D.; Fernandez-Madrid, I.J.; Power, M.D.; Lieberburg, I.; van Duinen, S.G.; Bots, G.T.; Luyendijk, W.; Frangione, B.
 Science 248, 1124-1126, 1990
 A;Title: Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type.
 A;Reference number: I39453; MUID:90260663; PMID:2111584
 A;Accession: I39453
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 656-737 <LEV>
 A;Cross-references: GB:M37896; NID:g178618; PIDN:AAA51727.1; PID:g178620
 A;Note: a mutation with 693-Gln is presented
 R;Murrell, J.; Farlow, M.; Ghetti, B.; Benson, M.D.
 Science 254, 97-99, 1991
 A;Title: A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease.
 A;Reference number: I59562; MUID:92022553; PMID:1925564
 A;Accession: I59562
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 689-716, 'F', 718-737 <MUR>
 A;Cross-references: GB:S57665; NID:g236720; PIDN:AAB19991.1; PID:g236721
 R;Kamino, K.; Orr, H.T.; Payami, H.; Wijsman, E.M.; Alonso, M.E.; Pulst, S.M.; Anderson, L.; O'dahl, S.; Nemens, E.; White, J.A.; Sadovnick, A.D.; Ball, M.J.; Kaye, J.; Warren, A.; McInnis, M.; Antonarakis, S.E.; Korenberg, J.R.; Sharma, V.; Kukull, W.; Larson, E.; Heston, L.L.; Martin, G.M.; Bird, T.D.; Schellenberg, G.D.
 Am. J. Hum. Genet. 51, 998-1014, 1992
 A;Title: Linkage and mutational analysis of familial Alzheimer disease kindreds for the APP gene region.
 A;Reference number: A44017; MUID:93035397; PMID:1415269
 A;Accession: A44017
 A;Molecule type: DNA
 A;Residues: 687-692, 'G', 694-718 <KAM1>
 A;Cross-references: GB:S45135; NID:g257377; PIDN:AAB23645.1; PID:g257378
 A;Experimental source: familial Alzheimer disease family SB
 A;Note: sequence extracted from NCBI backbone (NCBIP:115374)
 A;Accession: B44017
 A;Molecule type: DNA
 A;Residues: 687-718 <KAM2>
 A;Cross-references: GB:S45136; NID:g257379; PIDN:AAB23646.1; PID:g257380
 A;Experimental source: familial Alzheimer disease family LIT
 A;Note: sequence extracted from NCBI backbone (NCBIP:115376)
 A;Note: this sequence has a silent mutation
 R;Kang, J.; Lemaire, H.G.; Unterbeck, A.; Salbaum, J.M.; Masters, C.L.; Grzeschik, K.H.; Multhaup, G.; Beyreuther, K.; Muller-Hill, B.
 Nature 325, 733-736, 1987

A;Title: The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor.
 A;Reference number: A03134; MUID:87144572; PMID:2881207
 A;Accession: A03134
 A;Molecule type: mRNA
 A;Residues: 1-288,'V',365-770 <KAN>
 A;Cross-references: GB:Y00264; NID:g28525; PIDN:CAA68374.1; PID:g28526
 A;Note: alternative splice form APP(695)
 R;Robakis, N.K.; Ramakrishna, N.; Wolfe, G.; Wisniewski, H.M.
 Proc. Natl. Acad. Sci. U.S.A. 84, 4190-4194, 1987
 A;Title: Molecular cloning and characterization of a cDNA encoding the cerebrovascular and the neuritic plaque amyloid peptides.
 A;Reference number: A29030; MUID:87231971; PMID:3035574
 A;Accession: A29030
 A;Molecule type: mRNA
 A;Residues: 284-288,'V',365-646,'E',648-770 <ROB>
 A;Cross-references: GB:M16765; NID:g178539; PIDN:AAA51722.1; PID:g178540
 A;Note: the authors translated the codon GAG for residue 647 as Asp
 R;Goldgaber, D.; Lerman, M.I.; McBride, O.W.; Saffiotti, U.; Gajdusek, D.C.
 Science 235, 877-880, 1987
 A;Title: Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease.
 A;Reference number: A47584; MUID:87120328; PMID:3810169
 A;Accession: A47584
 A;Molecule type: mRNA
 A;Residues: 674-756,'S',758-770 <GOL>
 A;Cross-references: GB:M15533; NID:g178706; PIDN:AAA35540.1; PID:g178707
 A;Experimental source: brain
 R;Tanzi, R.E.; Gusella, J.F.; Watkins, P.C.; Bruns, G.A.P.; St George-Hyslop, P.; Van Keuren, M.L.; Patterson, D.; Pagan, S.; Kurnit, D.M.; Neve, R.L.
 Science 235, 880-884, 1987
 A;Title: Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus.
 A;Reference number: A47585; MUID:87120329; PMID:2949367
 A;Accession: A47585
 A;Molecule type: mRNA
 A;Residues: 674-703 <TAN1>
 A;Cross-references: GB:M15532; NID:g177957; PIDN:AAA51564.1; PID:g177958
 R;Dyrks, T.; Weidemann, A.; Multhaup, G.; Salbaum, J.M.; Lemaire, H.G.; Kang, J.; Mueller-Hill, B.; Masters, C.L.; Beyreuther, K.
 EMBO J. 7, 949-957, 1988
 A;Title: Identification, transmembrane orientation and biogenesis of the amyloid A4 precursor of Alzheimer's disease.
 A;Reference number: S02638; MUID:88296437; PMID:2900137
 A;Accession: S02638
 A;Molecule type: mRNA
 A;Residues: 672-678 <DYR>
 R;Tanzi, R.E.; McClatchey, A.I.; Lamperti, E.D.; Villa-Komaroff, L.; Gusella, J.F.; Neve, R.L.
 Nature 331, 528-530, 1988
 A;Title: Protease inhibitor domain encoded by an amyloid protein precursor mRNA associated with Alzheimer's disease.
 A;Reference number: S00707; MUID:88122640; PMID:2893290
 A;Accession: S00707
 A;Molecule type: mRNA
 A;Residues: 286-344,'I',365-366 <TAN2>
 A;Cross-references: EMBL:X06982; NID:g28817; PIDN:CAA30042.1; PID:g929612

A;Experimental source: promyelocytic leukemia cell line HL60
 A;Note: alternative splice form APP(751)
 R;Ponte, P.; Gonzalez-DeWhitt, P.; Schilling, J.; Miller, J.; Hsu, D.;
 Greenberg, B.; Davis, K.; Wallace, W.; Lieberburg, I.; Fuller, F.; Cordell, B.
 Nature 331, 525-527, 1988
 A;Title: A new A4 amyloid mRNA contains a domain homologous to serine proteinase
 inhibitors.
 A;Reference number: S00925; MUID:88122639; PMID:2893289
 A;Accession: S00925
 A;Molecule type: mRNA
 A;Residues: 1-344,'I',365-770 <PO2>
 A;Cross-references: GB:X06989; EMBL:Y00297; NID:g28720; PIDN:CAA30050.1;
 PID:g28721
 A;Note: alternative splice form APP(751)
 R;Kitaguchi, N.; Takahashi, Y.; Tokushima, Y.; Shiojiri, S.; Ito, H.
 Nature 331, 530-532, 1988
 A;Title: Novel precursor of Alzheimer's disease amyloid protein shows protease
 inhibitory activity.
 A;Reference number: A38949; MUID:88122641; PMID:2893291
 A;Accession: A38949
 A;Molecule type: mRNA
 A;Residues: 287-367 <KIT>
 A;Cross-references: GB:X06981; NID:g28816; PIDN:CAA30041.1; PID:g929611
 A;Experimental source: glioblastoma cell line
 A;Note: alternative splice form APP(770)
 R;Vitek, M.P.; Rasool, C.G.; de Sauvage, F.; Vitek, S.M.; Bartus, R.T.; Beer,
 B.; Ashton, R.A.; Macq, A.F.; Maloteaux, J.M.; Blume, A.J.; Octave, J.N.
 Brain Res. Mol. Brain Res. 4, 121-131, 1988
 A;Title: Absence of mutation in the beta-amyloid cDNAs cloned from the brains of
 three patients with sporadic Alzheimer's disease.
 A;Reference number: A30320
 A;Accession: A30320
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 284-288,'V',365-770 <VIT1>
 A;Accession: B30320
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 122-288,'V',365-770 <VIT2>
 A;Accession: C30320
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 606-770 <VIT3>
 R;Zain, S.B.; Salim, M.; Chou, W.G.; Sajdel-Sulkowska, E.M.; Majocha, R.E.;
 Marotta, C.A.
 Proc. Natl. Acad. Sci. U.S.A. 85, 929-933, 1988
 A;Title: Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer
 disease brain: coding and noncoding regions of the fetal precursor mRNA are
 expressed in the cortex.
 A;Reference number: A31087; MUID:88124954; PMID:2893379
 A;Accession: A31087
 A;Molecule type: mRNA
 A;Residues: 507-770 <ZAI>
 A;Cross-references: GB:M18734; NID:g178572; PIDN:AAA51726.1; PID:g178573
 A;Note: the authors translated the codon GAA for residue 599 as Gly, ACC for
 residue 603 as Val, GTG for residue 604 as Glu, GAG for residue 605 as Leu, CTT

Qy 721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 |||||
 Db 721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 2

A49795

Alzheimer's disease amyloid beta protein precursor - crab-eating macaque

C;Species: Macaca fascicularis (crab-eating macaque)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: A49795

R;Podlisny, M.B.; Tolan, D.R.; Selkoe, D.J.

Am. J. Pathol. 138, 1423-1435, 1991

A;Title: Homology of the amyloid beta protein precursor in monkey and human supports a primate model for beta amyloidosis in Alzheimer's disease.

A;Reference number: A49795; MUID:91273117; PMID:1905108

A;Accession: A49795

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-695 <POD>

A;Cross-references: GB:M58727; NID:g342062; PIDN:AAA36829.1; PID:g342063

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

C;Keywords: alternative splicing

Query Match 88.5%; Score 3590.5; DB 1; Length 695;
 Best Local Similarity 90.1%; Pred. No. 4.9e-185;
 Matches 694; Conservative 1; Mismatches 0; Indels 75; Gaps 1;

Qy 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
 |||||
 Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy 61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
 |||||
 Db 61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

Qy 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP CGIDKFR 180
 |||||
 Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP CGIDKFR 180

Qy 181 GVEFVCCPLAEESDNVDSADAEEEDSDVWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
 |||||
 Db 181 GVEFVCCPLAEESDNVDSADAEEEDSDVWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy 241 EADDDDEDDGDEVEEEAE EEPYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300
 |||||
 Db 241 EADDDDEDDGDEVEEEAE EEPYEEATERTTTSIATTTTTTTESVEEVVR----- 288

Qy 301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

 Db 289 ----- 288

Qy 361 PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEAAERQA 420
 :|||
 Db 289 ---VPTTAASTPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEAAERQA 345

Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAI	480
Db	346	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAI	405
Qy	481	QAVPPRPRHVFVNMLKKYVRAEQKDROHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	406	QAVPPRPRHVFVNMLKKYVRAEQKDROHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	465
Qy	541	MNQSLSLLYNVPAVAEEIQDEVDELLQKEQNYSDDVLNMISEPRISYGNDALMPSLTET	600
Db	466	MNQSLSLLYNVPAVAEEIQDEVDELLQKEQNYSDDVLNMISEPRISYGNDALMPSLTET	525
Qy	601	KTTVELLPVNGEFSLLDLQPWHSFGADSV PANTENEVEPV DARPAADRGLTTRPGSGLTN	660
Db	526	KTTVELLPVNGEFSLLDLQPWHSFGADSV PANTENEVEPV DARPAADRGLTTRPGSGLTN	585
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKL VFFAEDVGSNKGAI IGLMVGGVVIATVIVITL	720
Db	586	IKTEEISEVKMDAEFRHDSGYEVHHQKL VFFAEDVGSNKGAI IGLMVGGVVIATVIVITL	645
Qy	721	VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKM QQNGYENPTYKF FEQM QN	770
Db	646	VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKM QQNGYENPTYKF FEQM QN	695

S00550

N;Alternate names: beta-A4 amyloid protein

C;Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text change 13-Aug-1999

R; Shivers, B.D.; Hilbich, C.; Multhaup, G.; Salbaum, M.; Beyreuther, K.;

EMBO J. 7, 1365-1370, 1988

A;Reference number: S00550; MUID:88312583; PMID:2900758

A;Molecule type: mRNA

A;Cross-references: EMBL:X07648; NID:g55616; PIDN:CAA30488.1; PID:g55617

Science 241, 223-226, 1988

A;Reference number: A41245; MUID:88264430; PMID:2968652

A;Molecule type: protein

A;Note: evidence for heparan sulfate attachment

FEBS Lett. 349, 109-116, 1994

A;Reference number: S46251; MUID:94320627; PMID:7913895

Qy	541	MNQSLSLLYNVP	AAEEIQDEVDELLQKEQNYSD	DVLNMISEPRISYGNDALMP	SLTET	600
Db	466	MNQSLSLLYNVP	AAEEIQDEVDELLQKEQNYSD	DVLNMISEPRISYGNDALMP	SLTET	525
Qy	601	KTTVELLPVNGE	FSLDDLQPHWSFGADSV	PANTENEVEPVDARPAADR	GLTTRPGSGLTN	660
Db	526	KTTVELLPVNGE	FSLDDLQPHWPF	GVDSVPANTENEVEPVDARPAADR	GLTTRPGSGLTN	585
Qy	661	IKTEEISEVKMDA	EFRHDSGYEVHHQKL	VFFAEDVGSNKGAIIGLMVGGV	VIATVIVITL	720
Db	586	IKTEEISEVKMDA	EFHDSGFVRHQKL	VFFAEDVGSNKGAIIGLMVGGV	VIATVIVITL	645
Qy	721	VMLKKKQYTSI	HHGVVEVDAAVTPEERHLS	KMQQNGYENPTYKFFEQMQN		770
Db	646	VMLKKKQYTSI	HHGVVEVDAAVTPEERHLS	KMQQNGYENPTYKFFEQMQN		695

RESULT 4

A27485

Alzheimer's disease amyloid beta/A4 protein homolog precursor - mouse

N;Alternate names: proteinase nexin II

C;Species: Mus musculus (house mouse)

C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 13-Aug-1999

C;Accession: A27485; S19727; I49485

R;Yamada, T.; Sasaki, H.; Furuya, H.; Miyata, T.; Goto, I.; Sakaki, Y.

Biochem. Biophys. Res. Commun. 149, 665-671, 1987

A;Title: Complementary DNA for the mouse homolog of the human amyloid beta protein precursor.

A;Reference number: A27485; MUID:88106489; PMID:3322280

A;Accession: A27485

A;Molecule type: mRNA

A;Residues: 1-695 <YAM>

A;Cross-references: GB:M18373; NID:g191568; PIDN:AAA37139.1; PID:g309085

A;Experimental source: brain

R;de Strooper, B.; van Leuven, F.; van den Berghe, H.

Biochim. Biophys. Acta 1129, 141-143, 1991

A;Title: The amyloid beta protein precursor or proteinase nexin II from mouse is closer related to its human homolog than previously reported.

A;Reference number: S19727; MUID:92096458; PMID:1756177

A;Accession: S19727

A;Molecule type: mRNA

A;Residues: 1-210,'G',212-220,'S',222-396,'A',398-402,'T',404-448,'A',450-695 <STR>

A;Cross-references: EMBL:X59379

R;Izumi, R.; Yamada, T.; Yoshikai, S.; Sasaki, H.; Hattori, M.; Sakaki, Y.

Gene 112, 189-195, 1992

A;Title: Positive and negative regulatory elements for the expression of the Alzheimer's disease amyloid precursor-encoding gene in mouse.

A;Reference number: I49485; MUID:92209998; PMID:1555768

A;Accession: I49485

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-19 <RES>

A;Cross-references: GB:D10603; NID:g220328; PIDN:BAA01456.1; PID:g220329

C;Genetics:

A;Map position: 16C3

Db 646 VMLKKKQYTSIHHGVEVDAAVTPEERHLSKMOONGYENPTYKFFEMON 695

Db 351 AVDKYLENPNDENEHDFRLKAKERLEGKHREKMSSEVMKEWEEAEROAKNLPKADKKAVIO 410

Qy	434	HFQEKVESLEQEAA	NERQQLVETHMARVEAMLNDRRR	IALENYITALQAVPPRPRHVFNM	493
Db	411	HFQEKVESLEQEAA	KQRQQLVETHMARVEAMLNDRRR	IALENYITALQADPPRPRHVFNM	470
Qy	494	LKKYVRAEQKDRQHTLKHFEHVRM	VDPKKAAQIRSQVMTHLRVIYER	MNQSLSLLYNVPA	553
Db	471	LKKYVRAEQKDRQHTLKHFEHVRM	VDPKKAAQIRSQVMTHLRVINER	MNQSFSLLYKVPA	530
Qy	554	VAEEIQDEVDELLQKEQNYSD	DLANMISEPRISYGNDALMP	SLTETKTTVELLPVNGEF	613
Db	531	VAEEIQDEVDEL	FQKEQNYSDDMVSNMVS	DHRVSYGNDALMP	SLSETKTTVELLPVDGEF
					590
Qy	614	SLDDLQPWHSFGADSV	PANTENEVEPVDARPAADRGLT	TRPGSGLTNIKTEEISEVKMDA	673
Db	591	NIEDLQPWHSFGVDSV	PANTENEVEPVDARPAADRGLT	TRPGSGLTNIKTEEISEVKMDS	650
Qy	674	EFRHDSGYEVHHQKL	VFFAEDVGSNKGAIIGLMVGG	VVIATVIVITLVM	LKKKQYTSIHH
Db	651	EYRHDTAYEVHHQKL	VFFAEEVGSNKGAIIGLMVGG	VVIATVIVITLVM	LKKKQYTTIHH
Qy	734	GVVEVDAAVTPEERHLS	KMQQNGYENPTYKFFEQMQN		770
Db	711	GVVEVDAAVTPEERHLS	KMQQNGYENPTYKFFEQMQN		747

RESULT 6

A32761

hypothetical Alzheimer's disease amyloid beta protein, Alu-containing clone - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 29-Jan-1990 #sequence_revision 10-Apr-1996 #text_change 10-Apr-1996

C;Accession: A32761

R;de Sauvage, F.; Octave, J.N.

Science 245, 651-653, 1989

A;Title: A novel mRNA of the A4 amyloid precursor gene coding for a possibly secreted protein.

A;Reference number: A32761; MUID:89346754; PMID:2569763

A;Accession: A32761

A;Molecule type: mRNA

A;Residues: 1-484 <DES>

A;Cross-references: GB:M28373

A;Note: the authors translated the codon ATG for residue 433 as Leu

C;Comment: This is the hypothetical translation of a sequence believed to contain cloning artifacts.

C;Keywords: cloning artifact

Query Match 60.3%; Score 2446.5; DB 4; Length 484;
 Best Local Similarity 95.9%; Pred. No. 6.2e-124;
 Matches 463; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

Qy	80	LQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI	PYRCLVGEFVSDALLVPDKCKFLHQE	139
Db	1	LQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI	PYRCLVGEFVSDALLVPDKCKFLHQE	60
Qy	140	RMDVCETHLHWHTVAKETCSEKSTNLHDYGMLL	PCGIDKFRGVEFVCCPLAEESDNVDSA	199
Db	61	RMDVCETHLHWHTVAKETCSEKSTNLHDYGMLL	PCGIDKFRGVEFVCCPLAEESDNVDSA	120

Qy	200	DAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEEADDDDEDEDGDEVEEEEAE	259
Db	121	DAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEEADDDDEDEDGDEVEEEEAE	180
Qy	260	EPYEEATERTTSIATTTTTTTTTSVEEVVREVCSEQAETGPCRAMISRWFYDVTEGKCAPF	319
Db	181	EPYEEATERTTSIATTTTTTTTTSVEEVVREVCSEQAETGPCRAMISRWFYDVTEGKCAPF	240
Qy	320	FYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARDPVKLPPTTAASTPDAVDKYL	379
Db	241	FYGGCGGNRNNFDTEEYCMVCGSA-----IPTTAASTPDAVDKYL	281
Qy	380	ETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQAKNLPKADKKAVIQHFQEKV	439
Db	282	ETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQAKNLPKADKKAVIQHFQEKV	341
Qy	440	ESLEQEAAENERQQLVETHMARVEAMLNDRRRRLALENYITALQAVPPRPRHVFNMLKKYVR	499
Db	342	ESLEQEAAENERQQLVETHMARVEAMLNDRRRRLALENYITALQAVPPRPRHVFNMLKKYVR	401
Qy	500	AEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYERMNQSLSLLYNPAVAEEIQ	559
Db	402	AEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYERMNQSLSLLYNPAVAEEIQ	461
Qy	560	DEV	562
Db	462	DEV	464

RESULT 7

A49321

amyloid beta (A4) homolog 2 precursor - human

N;Alternate names: CDEI-binding protein

C;Species: Homo sapiens (man)

C;Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999

C;Accession: A49321; S34644; S40519

R;Sprecher, C.A.; Grant, F.J.; Grimm, G.; O'Hara, P.J.; Norris, F.; Norris, K.; Foster, D.C.

Biochemistry 32, 4481-4486, 1993

A;Title: Molecular cloning of the cDNA for a human amyloid precursor protein homolog: evidence for a multigene family.

A;Reference number: A49321; MUID:93250009; PMID:8485127

A;Accession: A49321

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-763 <SPR>

A;Cross-references: GB:S60099; NID:g300168; PIDN:AAC60589.1; PID:g300169

A;Experimental source: placenta

A;Note: sequence extracted from NCBI backbone (NCBIN:131198, NCBIIP:131199)

A;Note: expression was shown in placenta, brain, heart, lung, liver, and kidney

R;von der Kammer, H.; Klaudiny, J.; Hanes, J.; Scheit, K.H.

submitted to the EMBL Data Library, April 1993

A;Description: The human homologue of the murine CDEI binding protein is an amyloid precursor like protein.

A;Reference number: S34644

A;Accession: S34644

A;Molecule type: mRNA
A;Residues: 1-763 <VON>
A;Cross-references: EMBL:Z22572; NID:g394763; PIDN:CAA80295.1; PID:g394764
R;Wasco, W.; Gurubhagavatula, S.; Paradis, M.; Romano, D.M.; Sisodia, S.S.;
Hyman, B.T.; Neve, R.L.; Tanzi, R.E.
Nature Genet. 5, 95-99, 1993
A;Title: Isolation and characterization of APLP2 encoding a homologue of the
Alzheimer's associated amyloid beta protein precursor.
A;Reference number: S40519; MUID:94035131; PMID:8220435
A;Accession: S40519
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-763 <WAS>
A;Cross-references: GB:L27631; NID:g450391; PIDN:AAC41701.1; PID:g450392
C;Genetics:
A;Gene: GDB:APLP2; APPL2
A;Cross-references: GDB:139159; OMIM:104776
A;Map position: 11q23-11q25
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type
proteinase inhibitor homology
C;Keywords: alternative splicing; transmembrane protein
F;310-360/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>

Query Match 48.8%; Score 1981.5; DB 2; Length 763;
Best Local Similarity 50.4%; Pred. No. 9.2e-99;
Matches 408; Conservative 127; Mismatches 168; Indels 107; Gaps 21;

Qy	5	LALLLLAAWTARALEV-----PTDGNAG---LLAEPQIAMFCGRLNMHMNVQNGKWDSDP	56
		:	
Db	15	LLLLLLVGLTAPALALAGYIEALAAANAGTGFVAEPAEPQIAMFCGKLNMHVNIQTGKWEPPD	74
Qy	57	SGTKTCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIR	116
		: : :	
Db	75	TGTSKCFETKEEVLQYCQEMYPELQITNVMEANQVRSIDNWCRRDKKQCKS--RFVTPFK	132
Qy	117	CLVGEFVSDALLVDPKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMILLPCGI	176
		:	
Db	133	CLVGEFVSDVLLVPEKCFHFKERMEVCENHQHWHTVVKEACLTQGMTLYSYGMILLPCGV	192
Qy	177	DKFRGVEFVCCPLAEESDNVDSADAEEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAE	236
		: : : :	
Db	193	DQFHGTEYVCCPQTKIIGSVSKEEEEEDEE-----EEEEDEEEDYDVYKSEFPTEAD	245
Qy	237	VEE--EEA--DDDEDEDGDEVEEEAE-----EPYEEATERTTSIATTTTTTTTESVE	284
		: : : : : : :	
Db	246	LEDFTEAAVDEDEDEEEEGEEVVEDRDYDYDTFKGDDYNE--ENPTEPGSDGTMSDKEIT	303
Qy	285	EVVREVCSEQAETGPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNFPDTEEYCMVAVCGSA	344
		: : : :	
Db	304	HDVKAVCSQEAMTGPCRAVMPRWYFDLSKGGKCVRFIYGGCGGNRNFESEDYCMVAVCKAM	363
Qy	345	MSQSLLKTTQEPLARDPVKLPPTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRE	404
		: :	
Db	364	I-----PPTPLPT-----NDVDVYFETSADDNEHARFQKAKEQLEIRHRN	403
Qy	405	RMSQVMREWEEAERQAKNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAM	464
		: : : : : : : :	

Db 404 RMDRVKKEWEEAELQAKNLPKAERQTLIQHFQAMVKALEKEAASEKQQLVETHLARVEAM 463

Qy 465 LNDRRRLALENYITALQAVPPRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAA 524
 |||||:||||: |||: ||| : |::||| || ||::|::| |||:|||

Db 464 LNDRRRMALENYLAALQSDPPRPHRILQALRRYVRAENKDRLHTIRHYQHVLAVDPEKAA 523

Qy 525 QIRSQVMTHLRVIYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEP 584
 |::||| || || ||||| || |||:|::|:| |||:|::| |::|

Db 524 QMKSQVMTHLHVIEERNQSLSLLYKVPYVAQEIQEEIDELLQEQR-----ADM---- 572

Qy 585 RISYGNDALMPSLTETKTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARP 644
 | |::| | | | |::| | | | | | | | | | | | | | | | | |

Db 573 -----DQFTASISETPVDVR---VSSEES-EEIPPFHPF--HPFPALPENE----DTQP 616

Qy 645 AADRGLTTRPGSGLTN-----IKTEE---ISEVKMDAEFRHDSGYEVHHQKL VFFAEDVG 696
 : : |||: | | | | | | | | | | | | | | | | | | | | | |

Db 617 ELYHPM--KKGSGVGEQDGLIGAEKVINSKNKVDENMVIDETLDV--KEMIFNAERVG 672

Qy 697 S-----NKGAIIGLMVGGVVIATVIVITLVMLKKKQYTSIHHGVVEVD 739
 : |::|: | | |||||:| |||:| | | | | | | | | | | |

Db 673 GLEERESVGPLREDFSLSSALIGLLVIAVAIATVIVISLVMLRKRQYGTISHGIVEVD 732

Qy 740 AAVTPEERHLSKMQQNGYENPTYKFFEQMQ 769
 : |||||: ||| : |||||: |||

Db 733 PMLTPEERHLNKMQNHYENPTYKYLEQMQ 762

RESULT 8

S42880

amyloid precursor-like protein - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Mar-1999

C;Accession: S42880; S47528

R;Sandbrink, R.; Masters, C.L.; Beyreuther, K.

submitted to the EMBL Data Library, March 1994

A;Description: Complete nucleotide and deduced amino acid sequence of rat amyloid precursor-like protein 2 (Aplp2/Apph): Two amino acids length difference to human and murine homologues.

A;Reference number: S42880

A;Accession: S42880

A;Molecule type: mRNA

A;Residues: 1-765 <SAN>

A;Cross-references: EMBL:X77934

R;Sandbrink, R.; Masters, C.L.; Beyreuther, K.

Biochim. Biophys. Acta 1219, 167-170, 1994

A;Title: Complete nucleotide and deduced amino acid sequence of rat amyloid protein precursor-like protein 2 (APLP2/APPH): two amino acids length difference to human and murine homologues.

A;Reference number: S47528; MUID:94368849; PMID:8086458

A;Accession: S47528

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-765 <SA2>

A;Cross-references: EMBL:X77934

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

C;Keywords: alternative splicing

F;312-362/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>

Query Match 48.6%; Score 1971.5; DB 2; Length 765;
Best Local Similarity 49.7%; Pred. No. 3.2e-98;
Matches 401; Conservative 134; Mismatches 173; Indels 99; Gaps 20;

```
Qy      5 LALLLLAAWTARALEV-----PTDGNAG---LLAEPQIAMFCGRILNMHMNVQNGKWDSDP 56
      | :|||  || | :      |||  : ||||| ||||| : ||||| : | ||| : ||
Db     15 LLVLLLLLGLTAPAAALAGYIEALAANAGTGFVAEAPQIAMFCGKLNMHVNIQTGKWEPPD 74

Qy     57 SGTKTCTIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYR 116
      : ||| : | : ||| : ||||| : ||||| ||||| : ||||| | : ||| : ||| : : | ||| : :
Db     75 TGTKSCLTGKKEEVLQYCQEIYPELQITNVMEANQPVNIDSWCRRDKKQCRS--HIVIPFK 132

Qy    117 CLVGEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMILLPCGI 176
      ||||| ||||| : | : | ||||| : ||| | |||| | | : | : ||||| ||| :
Db    133 CLVGEFVSDVLLVPENCQFFHQERMEVCEKHQRWHTVVEACLTGEMTLYSYGMLLPCGV 192

Qy    177 DKFRGVEFVCCPLAE--ESDNVDSADAEEDSDVWVGADTDYA-DGSEDKVVEVAEEEE 233
      | : | | : |||| : : ||| : | : || : : : : ||| | || | :
Db    193 DQFHGTEYVCCPQTKVVDSDSTMSKEEEEEEEEE---DEEDYALDKSEFPTEADLEDFT 248

Qy    234 VAEVEEEEEADDEDEDEDGDEVVEEEAEPEYEE-----ATERTTSIATTTTTTTTSESVEEVV 287
      | : | : | : || : || : || : | : | : | : | : | : | : | : | : | :
Db    249 EAAADEDEDEEEEEEEEEEGEEVVEDRDYDDYDFSGKDDYNEENPTEPSSDGTISDKIEAHDV 308

Qy    288 REVCSEQAETGPCRAMISRWFYDVTGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQ 347
      : ||| : | ||||| : : ||||| : : ||| | ||||| ||||| : : ||||| : :
Db    309 KAVCSQEAMTGPCRAVMRWFYDLKSGKCVRFIYGGCGGNRNNFESEDYCMVCKTMI-- 366

Qy    348 SLLKTTQEPLARDPVKLPPTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMS 407
      | |||  : || | || | : |||| | ||||| : || : || ||
Db    367 -----PPTPLPT-----NDVDVYFETSADDNEHARFQKAKEQLEIRHRSRMD 408

Qy    408 QVMREWEAAERQAKNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLND 467
      : | : ||||| ||||| ||||| : : ||||| | : : || : || : || : || : || : ||
Db    409 RVKKEWEEAELQAKNLPKAERQTLIQHFQAMVKALEKEAAASEKQQLVETHLARVEAMLND 468

Qy    468 RRRIALENYITALQAVPPRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIR 527
      ||| : ||||| : ||| : ||| : | : ||||| ||| || : || : || ||| : ||| : :
Db    469 RRRIALENYLAALQSDPPRPHRILQALRRYVRAENKDRLHTIRHYQHVLAVDPEKAAQMK 528

Qy    528 SQVMTHLRVIYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRIS 587
      ||||| || || ||||| || ||| : ||| : ||| ||||| : : | : |
Db    529 SQVMTHLHVIEERRNQSLSLLYKVPYVAQEIQEEIDELLQEQR-----ADM----- 574

Qy    588 YGNDALMPSLTETKTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAAD 647
      | | : | | | : | : | : : | : | | : : ||| | : |
Db    575 ---DQFTSSISENPVDVR---VSSEES-EEIPPFHPF--HPFPSLSENE---DTQPELY 621

Qy    648 RGLTTRPGSGLTN-----IKTEE---ISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGS-- 697
      : : ||| : | || : | ||| | : | : : : | || ||
Db    622 HPM--KKGSGMAEQDGGGLIGAEKVINSKNKMNDENMVIDETLDV--KEMIFNAERVGGLE 677

Qy    698 -----NKGAIIGLMVGGVVIATVIVITLVMLKKKQYTSIHGHVVEVDAAV 742
      : | : ||| : | ||||| : ||||| : ||| : | ||| : |
Db    678 EEPDSVGPLREDFSLSSSALIGLLVIAVAIATVIVISLVMLRKRQYGTISHGIVEVHPML 737
```

Qy 743 TPEERHLSKMQQNGYENPTYKFPEQMQ 769
 |||||:|:|:|:|:|:|:|:|
 Db 738 TPEERHLNKMONGHYENPTYKYLEQMQ 764

RESULT 9

A49974

beta-amyloid precursor protein 2 homolog APLP2 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 06-Oct-1994 #sequence revision 18-Nov-1994 #text change 13-Aug-1999

C;Accession: A49974

R;Slunt, H.H.; Thinakaran, G.; Von Koch, C.; Lo, A.C.; Tanzi, R.E.; Sisodia, S.S.

J. Biol. Chem. 269, 2637-2644, 1994

A;Title: Expression of a ubiquitous, cross-reactive homologue of the mouse beta-amyloid precursor protein (APP).

A;Reference number: A49974; MUID:94132029; PMID:8300594

A;Accession: A49974

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: nucleic acid

A;Residues: 1-751 <SLU>

A;Cross-references: GB:U15571; NID:q558467; PIDN:AAA50603.1; PID:q558468

A;Note: sequence extracted from NCBI backbone (NCBIP:144636)

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

F;310-360/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>

Query Match 48.2%; Score 1956.5; DB 2; Length 751;
Best Local Similarity 49.3%; Pred. No. 2e-97;
Matches 397; Conservative 130; Mismatches 169; Indels 109; Gaps 20;

Qy	5	LALLLLAAWTARALEV----	PTDGNAG---	LIAEPQIAMFCGR	LNMMHNVQNGKWDSDP	56
		: :	:	:	:	
Db	15	LLVLLLLGLTAPAAALAGYI	EALANAGTGFAVAEPOI	AMFCGKLNMHVNI	OTGKWEPPD	74

Qy 57 SGTKTCIDTKEGILQYQCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYR 116
:||||:| | | :|||||:|||||||:||||| | :||:| ::|||: | |||::

Db 75 TGTKSCLGTKEEVLQYQCQEIYPELQITNVMEANQPVNIDSWCRRDKROCKS--HIVIPFK 132

Qy 117 CLVGEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGI 176
| | | | | | | | | | | | | : | | | | | : | | | | | : | | | : | | | : | | | | | :
Db 133 CLVGEFVSDVLLVPDNCQFFHQERMEVCEKHORWHTLVKEACLTGLTLYSYGMLLPCGV 192

Qy	177	DKFRGVEFVCCPLAE--ESDNVDSADAEEDDSDVWVGADTDYADGSEDKVVEVAE---E	231
		: : : : : : :: :: :	
Db	193	DQFHGTEYVCCPQTKTVDS DSTMSKEEEEEEE-----DEEDEFEEDYDLKSEFPTE	243

```

Qy      232 EEVAEVEEEEEAD-DEDEDGDGEVEEEAE-----EPYEEATERTTSIATTTTTTTT 282
          :: : | || ::||:|:|:| | : : : | | | | : :
Db      244 ADLEDFTAAADEEEEEDEEGEEVVEDRDYYYYDPFKGDYNE--ENPTPSSEGTISDKE 301

```

Qy 283 VEEVREVCSEQAETGPCRAMISRWFYDVTGKCAPFFYGGCGGNRNNFDTEEYCMVCG 342
: |:|::| |||||:: |||||::|| | ||||| |||||::|:| |||||
Db 302 IVHDKAVCSQEAMTGPCRAVMRWFYDLSKGKCVRFIYGGCGGNRNNFESEDCMAVCK 361

Qy 343 SAMSQSLLKTTQEPLARDPVKLPTTAASTPDAVDKYLETPGDENEHAHFOKAKERLEAKH 402

```

      : : | | | | : | | | | : | | | | | | | | : |
Db      362 AMI-----PPTPLPT-----NDVDVYFETSADDNEHARFQKAKEQLEIRH 401
Qy      403 RERMSQVMREWEEAERQAKNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVE 462
      | | : | : | | | | | | | | : : : | | | | | : | | | | | : | | | |
Db      402 RNRMDRVKKEWEEAELQAKNLPKTERQTLIQHFQAMVKALEKEAAASEKQQLVETHLARVE 461
Qy      463 AMLNDRRRRLALENYITALQAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPPK 522
      | | | | | | | : | | | | : | | | : | | | : | | | | | | | | | | |
Db      462 AMLNDRRRRIALENYLAALQSDPPRPHRILQALRRYVRAENKDLRLHTIRHYQHVLAVDPEK 521
Qy      523 AAQIRSQVMTHLRVIYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLVLANMIS 582
      | | | : : | | | | | | | | | | | | | | | | | | | | | : |
Db      522 AAQMKSQVMTHLHVIEERNQSLSLLYKVPYVAQEIQEEIDELLQEQR-----ADM-- 572
Qy      583 EPRI SYGNDALMPSLTETKTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDA 642
      | | | | : | | : | : | | | | | : | : | |
Db      573 -----DQFTSSISENPVDVRVSSEESE-EIPPFHPLHPF-----PSLSENE----- 612
Qy      643 RPAADRGLTTRPGSGLTNIKTEEI-SEVKMDAEFRHDSGYEVHHQKLVFFAEDVGS----- 697
      | : : | | | : : | | : | | | : | : : : | | | |
Db      613 ----GSGMAEQDG-GLIGAEKVINSKNKM DENMVIDETLDV--KEMIFNAERVGGLEEE 665
Qy      698 -----NKGAIIGLMVGGVVIATVIVITLVMLKKKQYTSIHGHGVVEVDAAVTP 744
      : | : | | : | | | | | | : | | | : | | : | | | | : |
Db      666 PESVGPLREDFSLSSNALIGLLVIAVAIATVIVISLVMLRKRQYGTISHGIVEVDPM LTP 725
Qy      745 EERHLSKMQQNGYENPTYKFFEQMQ 769
      | | | | : | | | : | | | | | : | | |
Db      726 EERHLNKMQN HGYENPTYKYLEQMQ 750

```

RESULT 10

A46362

amyloid precursor-like protein - mouse

C;Species: Mus musculus (house mouse)

C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999

C;Accession: A46362

R;Wasco, W.; Bupp, K.; Magendantz, M.; Gusella, J.F.; Tanzi, R.E.; Solomon, F.
Proc. Natl. Acad. Sci. U.S.A. 89, 10758-10762, 1992

A;Title: Identification of a mouse brain cDNA that encodes a protein related to
the Alzheimer disease-associated amyloid beta protein precursor.

A;Reference number: A46362; MUID:93066322; PMID:1279693

A;Accession: A46362

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1-653 <WAS>

A;Experimental source: brain

A;Note: sequence inconsistent with the nucleotide translation

A;Note: sequence extracted from NCBI backbone (NCBIN:118683, NCBIP:118684)

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type
proteinase inhibitor homology

C;Keywords: transmembrane protein

```

Query Match          28.4%;  Score 1150.5;  DB 2;  Length 653;
Best Local Similarity 35.4%;  Pred. No. 2.6e-54;
Matches 274;  Conservative 120;  Mismatches 228;  Indels 153;  Gaps 19;

```


C;Species: Mus musculus (house mouse)
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Feb-1997
 C;Accession: JC1404
 R;Vidal, F.; Blangy, A.; Rassoulzadegan, M.; Cuzin, F.
 Biochem. Biophys. Res. Commun. 189, 1336-1341, 1992
 A;Title: A murine sequence-specific DNA binding protein shows extensive local similarities to the amyloid precursor protein.
 A;Reference number: JC1404; MUID:93129193; PMID:1482349
 A;Accession: JC1404
 A;Molecule type: mRNA
 A;Residues: 1-511 <VID>
 C;Comment: This protein plays an important role in the early development of the mouse.
 C;Keywords: DNA binding; transmembrane protein

Query Match 27.4%; Score 1112.5; DB 2; Length 511;
 Best Local Similarity 43.0%; Pred. No. 2e-52;
 Matches 264; Conservative 87; Mismatches 136; Indels 127; Gaps 21;

Qy	174	CGIDKFRGVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEE	233
		: : : : : :	
Db	6	CGVDQFHGTEYVCCP---QTKTVDS-----DSTMSK-----EEEE	37
Qy	234	VAEVEEEEEADDDDEDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSE	293
		: : : : : : :	
Db	38	-----EEDEDEDEEEDYDLDKSEFPTEADLEDFTE---AAADEEEEEDEEEGEEVVED----	86
Qy	294	QAETGPCRAMISRWFYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTT	353
		: : : : : : : :	
Db	87	-----RDYYD-----PF---KGDDYNEENPTE-----PSSEGTI--S	114
Qy	354	QEPLARDPVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREW	413
		: : : : : : :	
Db	115	DKEIVHD-VKVPPTPLPTND-VDVYFETSADDNEHARFQKAKEQLEIRHRNRMDRVKKEW	172
Qy	414	EEAERQAKNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLAL	473
		: : : : : : :	
Db	173	EEAELQAKNLPKTERQTLIQHFQAMVKALEKEAAASEKQQLVETHLARVEAMLNDRRRRLAL	232
Qy	474	ENYITALQAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTH	533
		: : : : : : :	
Db	233	ENYLAALQSDPPRPHRILQALRRYVRAENKDRLHTIRHYQHVLAVDPEKAAQMKSQVMTH	292
Qy	534	LRVIYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRI SYGNDAL	593
		: : : : : : : :	
Db	293	LHVEEERRNQSLSLLYKVPYVAQEIQEEIDELLQEQR-----ADM-----DQF	335
Qy	594	MPSLTETKTTVELLPVNGEFSLLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTR	653
		: : : : : : : :	
Db	336	TSSISENPVDVRVSSESE-EIPPFHPLHPF-----PSLSENE-----GSGMAEQ	379
Qy	654	PGSGLTNIKTEEI-SEVKMDAEFRHDSGYEVHHQKLVFFAEDVGS-----	697
		: : : : : : :	
Db	380	DG-GLIGAEKVINSKNKMNDENMVIDETLDV--KEMIFNAERVGGLEEEEPESVGPLREDF	436
Qy	698	--NKGAIIGLMVGGVVIATVIVITLVMKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQN	755
		: : : : : : : : :	

Db 437 SLSSNALIGLLVIAVAIATVIVISLVMLRKRQYGTISHGIVEVDPMLTPEERHLNKMQNH 496
 QY 756 GYENPTYKFFEQMQ 769
 |||||: |||
 Db 497 GYENPTYKYLEQMQ 510

RESULT 12

T15795

hypothetical protein C42D8.8 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000

C;Accession: T15795; A49414

R;Hallsworth, K.

submitted to the EMBL Data Library, April 1996

A;Description: The sequence of *C. elegans* cosmid C42D8.

A;Reference number: Z18405

A;Accession: T15795

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-686 <HAL>

A;Cross-references: EMBL:U56966; NID:g1293844; PID:g1293850; PIDN:AAA98722.1; GSPDB:GN00028; CESP:C42D8.8

A;Experimental source: strain Bristol N2; clone C42D8

R;Daigle, I.; Li, C.

Proc. Natl. Acad. Sci. U.S.A. 90, 12045-12049, 1993

A;Title: apl-1, a *Caenorhabditis elegans* gene encoding a protein related to the human beta-amyloid protein precursor.

A;Reference number: A49414; MUID:94089766; PMID:8265668

A;Accession: A49414

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 7-686 <DAI>

A;Cross-references: GB:U00240; NID:g416296; PIDN:AAC46470.1; PID:g416297

C;Genetics:

A;Gene: CESP:C42D8.8

A;Map position: X

A;Introns: 22/3; 78/3; 121/1; 199/1; 230/1; 274/3; 344/3; 410/2; 471/2; 537/3; 580/3

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

Query Match 19.3%; Score 785; DB 2; Length 686;

Best Local Similarity 26.8%; Pred. No. 1e-34;

Matches 224; Conservative 111; Mismatches 272; Indels 230; Gaps 23;

QY 1 MLPGLIALLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHMNVQNGKWDSDPSGTK 60
 :: || : : | | | | | : | || | : | : | : |
 Db 6 LMIGLLIPILVA-TVYAEGSPAGSKRHEKFI PMVAFSCGYRNQYM-TEEGSWKTDDERYA 63
 QY 61 TCIDTKEGILQYCQEVYPQLQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVPIPYRCLVG 120
 || | ||: ||: || : |||: || : | : ||: || | | || | : |
 Db 64 TCFSGKLDILKYCRKAYPSMNITNIVEYSHEVSI SDWCREEGSPCK-WTHSVRPYHCIDG 122
 QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTN-----LHDYGMLLPC 174
 || | : || || | : | | | : || | : | | : : : : | ||
 Db 123 EFHSEALQVPHDCQFQSHVNSRDQCNDYQHWKDEAGKQCKTKKSKGNKMDIVRSFAVLEPC 182

Qy 175 GIDKFRGVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEV 234
 :| | ||||| :| :| :|
 Db 183 ALDMFTGVEFVCCP-----NDQTNKTDVQKTK----- 209

Qy 235 AEVEEEEADDEDEDGDEVEEEAEEPYEEATERTTSIATTTTTTTESVEEVVREVCSEQ 294
 |:|: ||||| |: ||:| ||
 Db 210 ---EEDDDDDDEDDAYEDDYSEESDEKDEE----- 236

Qy 295 AETGPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQ 354
 Db 237 ----- 236

Qy 355 EPLARDPVKLPPTTAASTPDAVDKYLET PGDENEHAHFQKAKERLEAKHRERMSQVMREWE 414
 || :|:| | :| || | :|:| :|:| ||:|:| :|:|:|
 Db 237 EPSSQDP-----YFKIANWTNEHDDFKKAEMRMDEKHKRKKVDKVMKEWG 280

Qy 415 EA-----ERQAKNLPKADKKAVIQ---HFQEKVESLEQEAAANERQQLVETHMARVEAMLN 466
 : |:|:|: || :| | |:| | ||:| | :|:| | |:|:|:|
 Db 281 DLETRYNEQKAKD-PKGAEKFKSQMNARFQKTVSSLEEEHKRMKEIEAVHEERVQAMLN 339

Qy 467 DRRRLALENYITAL--QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAA 524
 ::| | :| || | | | |:|:|:| || | :|:| | |:|:|
 Db 340 EKKRDATHDYRQALATHVNKPNKHSVLQSLKAYIRAEKDRMHTLNRYRHLLKADSKEAA 399

Qy 525 QIRSQVMTHLRVIYERMNQSLSLLYNVP-----AVA--EEIQDEVDELLQKEQNYS 573
 : |:|: || | |:|:|:| :| | | :|:|:| :|
 Db 400 AYKPTVIHRLRYIDLRLINGTLAMLRDFPDLEKYVRPIAVTYWKDYRDEVSPDISVE---- 455

Qy 574 DDVLANMISEPRISYGN--DALMPSLT----ETKTTVELLPVNGEFSLDDLQPWHSFGAD 627
 | | :| :| | | :| :| :|:| | :| :| :|
 Db 456 DSELTPIIHDEFKSKNAKLDVKAPTTTAKPVKETDNKVLPTASDSEEEADEYYEDED 515

Qy 628 SVPANT---ENEVEPVDARP-----AADRGLTTRPGSGLTNIKT 663
 | :|:|: || :| | | | | :| :| :|
 Db 516 EQVKKTPDMKKKVKVVDIKPKEIKVTIEEEKKAPKLVETSVQTDEDDDDDESSSSTSSES 575

Qy 664 EE-----ISEVKMDAE-----FRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGG 709
 :| | |:|:| | :|| | :| | | :| :| :|
 Db 576 DEDEDKNIKELRVDIEPIIDEPASFYRHD-----KLIQSPEVERSASSVFQPYVLAS 627

Qy 710 VVIATVIVITLVMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFE 766
 : | | | : : | :||| |||||:| | ||||| ||:|
 Db 628 AMFITAICIIAFAITNARRRRAMRGFIEVD-VYTPEERHVAGMQVNGYENPTYSF 683

RESULT 13

A32758

beta-amyloid-like protein precursor - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 08-Dec-1989 #sequence_revision 08-Dec-1989 #text_change 24-Sep-1998

C;Accession: A32758

R;Rosen, D.R.; Martin-Morris, L.; Luo, L.; White, K.

Proc. Natl. Acad. Sci. U.S.A. 86, 2478-2482, 1989

A;Title: A *Drosophila* gene encoding a protein resembling the human beta-amyloid protein precursor.

A;Reference number: A32758; MUID:89184650; PMID:2494667

A;Accession: A32758
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-886 <ROS>
 A;Cross-references: GB:J04516; NID:g158371; PID:g158372
 C;Genetics:
 A;Gene: FlyBase:Appl
 A;Cross-references: FlyBase:FBgn0000108
 C;Keywords: transmembrane protein

Query Match 18.2%; Score 737.5; DB 2; Length 886;
 Best Local Similarity 24.7%; Pred. No. 5.1e-32;
 Matches 235; Conservative 137; Mismatches 313; Indels 265; Gaps 30;

Qy	7	LLLLAAWTARALEVPTDGNAGLLA-----EPQIAMFC--GRLNMHMNV-QNGKWDSDPSG	58
		: : : : : : :	
Db	9	LLLRSLWVVLAI-----GTAQVQAASPRWEPQIAVLCEAGQIYQPQYLSEEGRWVTDLSK	63
Qy	59	T---KTCIDTKEGILQYCQEVPELQITNVVEANQPVTIQNWCKRG---RKQCKTHPHFV	112
		: : : : : : : :	
Db	64	KTTGPTCLRDKMDLLDYCKKAYPNRDI TNIVESSHYQKIGGWCROGALNAACKGSHRWI	123
Qy	113	IPYRCLVGEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMML	172
		: : : : : :	
Db	124	KPFRCL-GPFQSDALLVPEGCLFDHIHNASRCWPFVRWNQTGAAACQERGMQMRTFAMLL	182
Qy	173	PCGIDKFRGVFVCCP-----LAEESDNVD---SA	199
		: : :	
Db	183	PCGISVFSGVFVCCPKHFKTDEIHVKKTDLPVMPAAQINSANDELMNDEDDSDNSNYSK	242
Qy	200	DAEEDDSVWVGADTDYADGSEDKVVEVAEEEEV-----AEV	237
		: : : :	
Db	243	DANEDDL-----DEDDLMGDDEEDDMVADEAATAGGSPNTGSSGDSNSGSLDDINA EY	296
Qy	238	EE-EEADDDDEDEDGDEVEEEAE EPYEEA-TERTTSIATTTTTTTESVEEVVREVCSEQA	295
		: : : : : : : : : : : :	
Db	297	DSGEEGDNYEEDGAGSESEAEVEASWDQSGGAKVVS LKSDSSSPSSAPVAPAPEKAPVKS	356
Qy	296	ETGPCRAMISRWFYDVTEGKCAPFFYGGCGGNRN NFDTEEYCMVCGSAMSQSLLKTTQE	355
		: : : : : : : : :	
Db	357	ESVTSTPQLS-----ASAAAFVAANSNGSGT-----GAGAPPSTAQPTSD	396
Qy	356	PLARDPVKLPPTTAASTPDVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEE	415
		: : : : : : : : :	
Db	397	P-----YFTHFDPHYEHQSYKVSQKRLEESHREKVTRVMKDWS	435
Qy	416	AERQAKNLPKADKKA-----VIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDR	468
		: : : : : : : : :	
Db	436	LEEKYQDMRLADPKAAQSFQKQRM TARFQTSVQALEEEGNAEKHQLAAMHQQRVLAHINQR	495
Qy	469	RRLALENYITALQAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEH-VRMVDP---KKA	524
		: : : : : :	
Db	496	KREAMTCYTQALTEQPPNAHHVEKCLQKLLRALHKDRAHALAHYRHLNLSGGPGGLEAAA	555
Qy	525	QIRSQVMTHLRVIYERMNQSLSLLYNVPAAVEEI-----QDEV-----	562
		: : : : : : : : :	
Db	556	SERPRTLRLIDIDRAVNQSM TMLKRYPELSAKIAQLMNDYILALRSKDDIPGSSIGMSE	615

```

Qy      563 -----DELLQKEQNYSDVLAN 579
          :| | | | | | | |
Db      616 EAEAGILDKYRVEIERKVAEKERLRLAEKQRKEQRAAEREKLRLEAKKVDDMLKS 675

Qy      580 MISE-----PRISYGNDA-----PSLTETKTTVELLPVNG 611
          :| | | | | | | |
Db      676 QVAEQSQPTQSSTQSQAQQQQQEKSLPGKELGPDAALVTAANPNLETTKS----- 726

Qy      612 EFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTNIKTEEISEVKM 671
          | | | | : | : | | | | : : :
Db      727 EKDLSDTE----YGEATVSTTKVQTVLP TVDDDAVQRAVEDVAAA-----VAHQEA 773

Qy      672 DAEFRHDSGYEVHHQKLVF-----FAEDVGSNK---GAIIGLMVGGVVIATVIVITLVML 723
          : : : | : : | : | | : | : : :
Db      774 EPQVQHFMTDHLGHRSSFSRLRREFAQHAAAKEGRNVYFTLSFAGIALMAAVFVGVA 833

Qy      724 KKKQYTSIH-HGVVEVDAAVTP-----EERHLSKMQQNGYENPTYKFFE 766
          | : | | : | | | | | : : | | | | | : |
Db      834 KWRTSRSPHAQGFIEVDQNVTHHP I VREEKIVPNMQINGYENPTYKYFE 883

```

RESULT 14

S38344

CDEI-binding protein - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 03-May-1996

C;Accession: S38344

R;Hanes, J.; von der Kammer, H.; Kristjansson, G.I.; Scheit, K.H.

Biochim. Biophys. Acta 1216, 154-156, 1993

A;Title: The complete cDNA coding sequence for the mouse CDEI binding protein.

A;Reference number: S38344; MUID:94032480; PMID:8218408

A;Accession: S38344

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-246 <HAN>

A;Cross-references: EMBL:Z22592

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

Query Match 17.4%; Score 706; DB 2; Length 246;
 Best Local Similarity 51.5%; Pred. No. 4.9e-31;
 Matches 136; Conservative 35; Mismatches 51; Indels 42; Gaps 7;

```

Qy      5 LALLLLAAWTARALEV-----PTDGNAG---LLAEPQIAMFCGRNLNMHMNVQNGKWDSDP 56
          | : | | | | | | : | | | | | | | | : | | | | |
Db      15 LLVLLLLLGLTAPAAALAGYIEALANAGTGF A VAE P Q I A M L C G K L N M H V N I Q T G K W E P D P 74

Qy      57 SGTKTCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYR 116
          : | | | : | | | : | | | | : | | | | | : | | | : | | | : |
Db      75 TGTKSCLGTKEEVLQYCQEIYPELQITNVMEANQPVNIDSWCRRDKRQCKS--HIVIPFK 132

Qy      117 CLVGEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGI 176
          | | | | | | | | | | | | : | | | | : | | | | | | : | | | | | :
Db      133 CLVGEFVSDVLLVPDNCQFFHQERMEVCEKHQRWHTLVKEACLTEGLTLYSYGMLLPCGV 192

Qy      177 DKFRGVEFVCCPLAEESDNVDSADAEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAE 236

```

```

      | : | | | : || | | : : | | |
Db      193 DQFHGTEYVCCP---QTKTVDS-----DSTMSKEEEEE--- 222

Qy      237 VEEEEADDDDED-DEDGDEVEEEAE 259
      ||:| |::|| | | | | | :
Db      223 -EEDEEDEEEDYDLDKSEFPTEAD 245

```

RESULT 15

A32282

Alzheimer's disease amyloid beta protein precursor - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Aug-1989 #sequence_revision 17-Aug-1989 #text_change 13-Aug-1999

C;Accession: A32282

R;Yamada, T.; Sasaki, H.; Dohura, K.; Goto, I.; Sakaki, Y.

Biochem. Biophys. Res. Commun. 158, 906-912, 1989

A;Title: Structure and expression of the alternatively-spliced forms of mRNA for the mouse homolog of alzheimer's disease amyloid beta protein precursor.

A;Reference number: A32282; MUID:89149813; PMID:2493250

A;Accession: A32282

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-100 <YAM>

A;Cross-references: GB:M24397; NID:g200350; PIDN:AAA39929.1; PID:g200351

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

C;Keywords: alternative splicing

F;11-61/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>

```

Query Match          12.3%;  Score 501;  DB 2;  Length 100;
Best Local Similarity 92.0%;  Pred. No. 1.5e-20;
Matches 92;  Conservative 2;  Mismatches 6;  Indels 0;  Gaps 0;

```

```

Qy      281 ESVEEVVREVCSEQAETGPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCM 340
      |||||
Db      1  ESVEEVVREVCSEQAETGPCRAMISRWFVDVTEGKCVPPFFYGGCGGNRNNFDTEEYCM 60

Qy      341 CGSAMSQSLKTTQEPLARDPVKLPPTTAASTPDAVDKYLE 380
      ||| :|||
Db      61 CGSVSTQSLKTTSEPLPQDPDKLPPTTAASTPDAVDKYLE 100

```

Search completed: December 13, 2003, 13:30:51

Job time : 51.9487 secs

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:30:00 ; Search time 87.4359 Seconds
 (without alignments)
 1637.856 Million cell updates/sec

Title: US-09-785-215-2
 Perfect score: 4058
 Sequence: 1 MLPGLALLLLAAWTARALEV.....KMQQNGYENPTYKFFEQMQN 770

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
 Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Match	Length	DB	ID	Description
No.		%				
		Query				

1	4058	100.0	770	9	US-09-794-927-55	Sequence 55, Appl
2	4058	100.0	770	9	US-09-795-847-55	Sequence 55, Appl
3	4058	100.0	770	9	US-09-794-743-55	Sequence 55, Appl
4	4058	100.0	770	9	US-09-794-748-55	Sequence 55, Appl
5	4058	100.0	770	9	US-09-904-987-2	Sequence 2, Appli
6	4058	100.0	770	9	US-09-794-925-55	Sequence 55, Appl
7	4058	100.0	770	9	US-09-681-442-55	Sequence 55, Appl
8	4058	100.0	770	10	US-09-149-718-6	Sequence 6, Appli
9	4058	100.0	770	10	US-09-785-215-2	Sequence 2, Appli
10	4058	100.0	770	11	US-09-848-616-172	Sequence 172, App
11	4058	100.0	770	11	US-09-869-414-55	Sequence 55, Appl
12	4058	100.0	770	11	US-09-548-366-55	Sequence 55, Appl
13	4058	100.0	770	12	US-10-335-035-3	Sequence 3, Appli
14	4058	100.0	770	12	US-10-223-809A-2	Sequence 2, Appli
15	4058	100.0	770	12	US-10-010-942B-38	Sequence 38, Appl
16	4058	100.0	770	12	US-10-357-935-3	Sequence 3, Appli
17	4058	100.0	770	12	US-10-050-898-218	Sequence 218, App
18	4058	100.0	770	12	US-10-050-902-218	Sequence 218, App
19	4058	100.0	770	12	US-10-427-208-75	Sequence 75, Appl
20	4058	100.0	770	15	US-10-217-584-5	Sequence 5, Appli
21	4058	100.0	770	15	US-10-204-362-2	Sequence 2, Appli
22	4058	100.0	770	15	US-10-169-580-5	Sequence 5, Appli
23	4058	100.0	772	9	US-09-794-927-59	Sequence 59, Appl
24	4058	100.0	772	9	US-09-795-847-59	Sequence 59, Appl
25	4058	100.0	772	9	US-09-794-743-59	Sequence 59, Appl
26	4058	100.0	772	9	US-09-794-748-59	Sequence 59, Appl
27	4058	100.0	772	9	US-09-794-925-59	Sequence 59, Appl
28	4058	100.0	772	9	US-09-681-442-59	Sequence 59, Appl
29	4058	100.0	772	11	US-09-869-414-59	Sequence 59, Appl
30	4058	100.0	772	11	US-09-548-366-59	Sequence 59, Appl
31	3943.5	97.2	751	9	US-09-794-927-57	Sequence 57, Appl
32	3943.5	97.2	751	9	US-09-795-847-57	Sequence 57, Appl
33	3943.5	97.2	751	9	US-09-794-743-57	Sequence 57, Appl
34	3943.5	97.2	751	9	US-09-794-748-57	Sequence 57, Appl
35	3943.5	97.2	751	9	US-09-794-925-57	Sequence 57, Appl
36	3943.5	97.2	751	9	US-09-681-442-57	Sequence 57, Appl
37	3943.5	97.2	751	10	US-09-149-718-4	Sequence 4, Appli
38	3943.5	97.2	751	11	US-09-869-414-57	Sequence 57, Appl
39	3943.5	97.2	751	11	US-09-548-366-57	Sequence 57, Appl
40	3943.5	97.2	751	12	US-10-357-935-2	Sequence 2, Appli
41	3943.5	97.2	751	12	US-10-427-208-74	Sequence 74, Appl
42	3943.5	97.2	751	15	US-10-169-580-4	Sequence 4, Appli
43	3943.5	97.2	753	9	US-09-794-927-61	Sequence 61, Appl
44	3943.5	97.2	753	9	US-09-795-847-61	Sequence 61, Appl
45	3943.5	97.2	753	9	US-09-794-743-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1

US-09-794-927-55

; Sequence 55, Application US/09794927

; Patent No. US20010016324A1

; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.


```

; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR,
AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/794,927
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-927-55

```

```

Query Match          100.0%; Score 4058; DB 9; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.2e-257;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
      |||
Db     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

Qy    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241 EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241 EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300

Qy    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

```

Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDAIMPSTTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDAIMPSTTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMN	770
Db	721	VMLKKKQYTSIIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMN	770

RESULT 2

US-09-795-847-55

; Sequence 55, Application US/09795847

; Patent No. US20010018208A1

; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Bienkowski, Michael J.

; APPLICANT: Heinrikson, Robert L.

; APPLICANT: Parodi, Luis A.

; APPLICANT: Yan, Riqiang

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND

; TITLE OF INVENTION: USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: 28341/6280DE

; CURRENT APPLICATION NUMBER: US/09/795,847

; CURRENT FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: 09/416,901

; PRIOR FILING DATE: 1999-10-13

; PRIOR APPLICATION NUMBER: 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-847-55
```

Query Match 100.0%; Score 4058; DB 9; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.2e-257;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG	120
Db	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEDDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEDDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE	240
Qy	241	EADDDDEDEDGDEVEEEAEOPYEEATERTTSIATTTTTTTTESVEEVVREVCSEAETGPC	300
Db	241	EADDDDEDEDGDEVEEEAEOPYEEATERTTSIATTTTTTTTESVEEVVREVCSEAETGPC	300
Qy	301	RAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNLALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNLALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLLDLQPWHSFGADSVANTENEVEPVDPARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLLDLQPWHSFGADSVANTENEVEPVDPARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720

```

Db          661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA TVIVITL 720
Qy          721 VMLKKKQYTSIHG VVEVDAAVTPEERHLSKMQQNGYENPTYKFFE QMQN 770
          |||
Db          721 VMLKKKQYTSIHG VVEVDAAVTPEERHLSKMQQNGYENPTYKFFE QMQN 770

```

RESULT 3

US-09-794-743-55

```

; Sequence 55, Application US/09794743
; Patent No. US20010021391A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR,
AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280BC
; CURRENT APPLICATION NUMBER: US/09/794,743
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-743-55

```

```

Query Match          100.0%; Score 4058; DB 9; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.2e-257;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy          1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGR LNMHMNVQNGKWDSDPSG TK 60
          |||
Db          1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGR LNMHMNVQNGKWDSDPSG TK 60

Qy          61  TCIDTKEGILQYCQEVYP ELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
          |||
Db          61  TCIDTKEGILQYCQEVYP ELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

Qy          121 EFVSDALLVDPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGM LLLPCGIDKFR 180
          |||
Db          121 EFVSDALLVDPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGM LLLPCGIDKFR 180

```



```

; CURRENT APPLICATION NUMBER: US/09/794,748
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-748-55

```

```

Query Match          100.0%; Score 4058; DB 9; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.2e-257;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
      |||
Db     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

Qy    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP CGIDKFR 180
      |||
Db    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP CGIDKFR 180

Qy    181  GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181  GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241  EADDDDEDDDEDGDEVEEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241  EADDDDEDDDEDGDEVEEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300

Qy    301  RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301  RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361  PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420
      |||
Db    361  PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420

Qy    421  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480
      |||
Db    421  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480

Qy    481  QAVPPRPRHVFNM LKKYVRAEQKDRQHTLKHFEHVRMVDPKKAQIRSQVMTHLRVIYER 540

```

Db	481		QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541		MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET	600
Db	541		MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET	600
Qy	601		KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601		KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661		IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661		IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721		VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770
Db	721		VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770

RESULT 5

US-09-904-987-2

; Sequence 2, Application US/09904987

; Patent No. US20020037908A1

; GENERAL INFORMATION:

; APPLICANT: No. US20020037908A1actyl, Inc.

; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prepathological

; TITLE OF INVENTION: Protein Assembly or Aggregation

; FILE REFERENCE: 42108/26146

; CURRENT APPLICATION NUMBER: US/09/904,987

; CURRENT FILING DATE: 2001-07-12

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 770

; TYPE: PRT

; ORGANISM: homo sapiens

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: NCBI ENTREZ / QRHUA4

; DATABASE ENTRY DATE: 2000-09-15

; RELEVANT RESIDUES: (1)..(770)

US-09-904-987-2

Query Match 100.0%; Score 4058; DB 9; Length 770;

Best Local Similarity 100.0%; Pred. No. 1.2e-257;

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIYRCLVG	120
Db	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180

Db	121		180
Qy	181	GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Db	181		240
Qy	241	EADDDDEDEDGDEVEEEAEPPYEEATERTTTSIATTTTTTTTESVEEVVREVCSEAETGPC	300
Db	241		300
Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301		360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Db	361		420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	421		480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481		540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNDAIMPSTET	600
Db	541		600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601		660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661		720
Qy	721	VMLKKKQYTSIHGCVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEOMQN	770
Db	721		770

RESULT 6

US-09-794-925-55

; Sequence 55, Application US/09794925

; Patent No. US20020064819A1

; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Bienkowski, Michael J.

; APPLICANT: Heinrikson, Robert L.

; APPLICANT: Parodi, Luis A.

; APPLICANT: Yan, Riqiang

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES

; TITLE OF INVENTION: THEREFOR


```

; FILE REFERENCE: 28341/6280HI
; CURRENT APPLICATION NUMBER: US/09/794,925
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-925-55

```

```

Query Match          100.0%; Score 4058; DB 9; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.2e-257;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHMNVQNGKWDSDPSGTK 60

Qy     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
      |||
Db     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

Qy    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181  GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181  GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241  EADDDDEDDEDGDEVEEEAEPEYEEATERTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241  EADDDDEDDEDGDEVEEEAEPEYEEATERTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300

Qy    301  RAMISRWYFDVTEGKCAPPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301  RAMISRWYFDVTEGKCAPPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361  PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420
      |||
Db    361  PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420

Qy    421  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480
      |||
Db    421  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480

```

Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRISYGN DALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRISYGN DALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPV DARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPV DARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770
Db	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770

RESULT 7

US-09-681-442-55

; Sequence 55, Application US/09681442

; Patent No. US20020081634A1

; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Bienkowski, Michael J.

; APPLICANT: Heinrikson, Robert L.

; APPLICANT: Parodi, Luis A.

; APPLICANT: Yan, Riqiang

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: 28341/6280FG

; CURRENT APPLICATION NUMBER: US/09/681,442

; CURRENT FILING DATE: 2001-04-05

; PRIOR APPLICATION NUMBER: 09/416,901

; PRIOR FILING DATE: 1999-10-13

; PRIOR APPLICATION NUMBER: 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 55

; LENGTH: 770

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-681-442-55

Query Match	100.0%;	Score 4058;	DB 9;	Length 770;
Best Local Similarity	100.0%;	Pred. No. 1.2e-257;		

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDPSGTK 60
      |||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDPSGTK 60

Qy     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
      |||
Db     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

Qy    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181 GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE 240
      |||
Db    181 GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE 240

Qy    241 EADDDDEDDEDGDEVEEEAEEP YEEATERTTSIATTTTTTTTESVEEVVREVCSEAETGPC 300
      |||
Db    241 EADDDDEDDEDGDEVEEEAEEP YEEATERTTSIATTTTTTTTESVEEVVREVCSEAETGPC 300

Qy    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361 PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
      |||
Db    361 PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420

Qy    421 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480
      |||
Db    421 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480

Qy    481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
      |||
Db    481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Qy    541 MNQSLSLLYNVPAAVAEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNDALMPSLTET 600
      |||
Db    541 MNQSLSLLYNVPAAVAEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNDALMPSLTET 600

Qy    601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
      |||
Db    601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Qy    661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
      |||
Db    661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720

Qy    721 VMLKKKQYTSIHGCVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMQN 770
      |||
Db    721 VMLKKKQYTSIHGCVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMQN 770

```

```

; Sequence 6, Application US/09149718
; Patent No. US20020104104A1
; GENERAL INFORMATION:
;   APPLICANT: Dora K. Games, Dale B. Schenk, Lisa C. McConlogue,
;   APPLICANT: Peter A. Seubert, and Russell E. Rydel
;   TITLE OF INVENTION: Method For Identifying Alzheimer's Disease
;   TITLE OF INVENTION: Therapeutics Using Transgenic Animal Models
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Patrea L. Pabst
;     STREET: 2800 One Atlantic Center
;     STREET: 1201 West Peachtree Street
;     CITY: Atlanta
;     STATE: GA
;     COUNTRY: USA
;     ZIP: 30309-3450
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/149,718
;     FILING DATE:
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/08/660,487
;     FILING DATE:
;     APPLICATION NUMBER: 08/480,653
;     FILING DATE: June 7, 1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Pabst, Patrea L.
;     REGISTRATION NUMBER: 31,284
;     REFERENCE/DOCKET NUMBER: ANS101CIP
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (404)-873-8794
;     TELEFAX: (404)-873-8795
;   INFORMATION FOR SEQ ID NO: 6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 770 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-09-149-718-6

```

```

Query Match          100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.2e-257;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MLPGLALLLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MLPGLALLLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

```

```

Qy      61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

```

Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Qy	241	EADDDDEDEDGDEVEEEAEEPYEEATERTTTSIATTTTTTTTSESVEEVVREVCSEQAETGPC	300
Db	241	EADDDDEDEDGDEVEEEAEEPYEEATERTTTSIATTTTTTTTSESVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPVAEELIQDEVDLLQKEQNYSDDLANMISEPRI SYGNDALMPSLTET	600
Db	541	MNQSLSLLYNVPVAEELIQDEVDLLQKEQNYSDDLANMISEPRI SYGNDALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAI IGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAI IGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFPEQM QN	770
Db	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFPEQM QN	770

RESULT 9

US-09-785-215-2

; Sequence 2, Application US/09785215

; Publication No. US20020187157A1

; GENERAL INFORMATION:

; APPLICANT: JENSEN, Martin Roland et al.

; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID

; FILE REFERENCE: 3631-0107P

; CURRENT APPLICATION NUMBER: US/09/785,215

; CURRENT FILING DATE: 2001-02-20

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn version 3.1

Query Match		100.0%;	Score 4058;	DB 10;	Length 770;
Best Local Similarity		100.0%;	Pred. No. 1.2e-257;		
Matches 770;		Conservative	0;	Mismatches	0; Indels 0; Gaps 0;

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGR_LNMHMNVQN_KWSDSPSGTK	60
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGR_LNMHMNVQN_KWSDSPSGTK	60
Qy	61	TCIDTKEGILQYCQEVPPELQITNVVEANQPVTIQNWCKRGRK_QCKTHPHFVIPYRCLVG	120
Db	61	TCIDTKEGILQYCQEVPPELQITNVVEANQPVTIQNWCKRGRK_QCKTHPHFVIPYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP_CGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP_CGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGS_EDKVVEVAEEEEVAEVEEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGS_EDKVVEVAEEEEVAEVEEEE	240
Qy	241	EADDDDEDEDGDGEVEEEEAEPYEEATERTTTSIATTTTTTTES_VEEVVREVCSEAETGPC	300
Db	241	EADDDDEDEDGDGEVEEEEAEPYEEATERTTTSIATTTTTTTES_VEEVVREVCSEAETGPC	300
Qy	301	RAMISRWFVDVTEGKCAPPFYGGCGGNRN_NFDTEEYCMAVCGSAMSQSLLKTQTQEPLARD	360
Db	301	RAMISRWFVDVTEGKCAPPFYGGCGGNRN_NFDTEEYCMAVCGSAMSQSLLKTQTQEPLARD	360
Qy	361	PVKLPPTTAASTPDADV_KYLETPGDENEHAHFQKAKERLEAKHR_RMSQVMREWEEAERQA	420
Db	361	PVKLPPTTAASTPDADV_KYLETPGDENEHAHFQKAKERLEAKHR_RMSQVMREWEEAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQE_AANERQQLVETHMARVEAMLNDRRR_LALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQE_AANERQQLVETHMARVEAMLNDRRR_LALENYITAL	480
Qy	481	QAVPPRPRHVF_NMLKKYVRAEQKDRQH_TLKHFEHVRMVDPKKAAQIRSQVMTHLRVI_YER	540
Db	481	QAVPPRPRHVF_NMLKKYVRAEQKDRQH_TLKHFEHVRMVDPKKAAQIRSQVMTHLRVI_YER	540
Qy	541	MNQSL_SLLYNVPAVAEEIQDEVDELLQKEQNYSDDVLANMISEPRI_SYGNDALMP_SLTET	600
Db	541	MNQSL_SLLYNVPAVAEEIQDEVDELLQKEQNYSDDVLANMISEPRI_SYGNDALMP_SLTET	600
Qy	601	KTTVELLPVN_GEFSDDLQP_WHSFGADSV_PANTENEVEPV_DARPAADRGLTTRPG_SGLTN	660
Db	601	KTTVELLPVN_GEFSDDLQP_WHSFGADSV_PANTENEVEPV_DARPAADRGLTTRPG_SGLTN	660
Qy	661	IKTEEISEVKMDAEFR_HDSGYEVHHQKL_VFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFR_HDSGYEVHHQKL_VFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720

Qy 721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 |||||
 Db 721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 10

US-09-848-616-172

; Sequence 172, Application US/09848616
 ; Publication No. US20030054010A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sebbel, Peter
 ; APPLICANT: Dunant, Nicolas
 ; APPLICANT: Bachmann, Martin
 ; APPLICANT: Tissot, Alain
 ; APPLICANT: Lechner, Franziska
 ; TITLE OF INVENTION: Molecular Antigen Array
 ; FILE REFERENCE: 1700.0180002
 ; CURRENT APPLICATION NUMBER: US/09/848,616
 ; CURRENT FILING DATE: 2001-05-05
 ; NUMBER OF SEQ ID NOS: 186
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 172
 ; LENGTH: 770
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-848-616-172

Query Match 100.0%; Score 4058; DB 11; Length 770;
 Best Local Similarity 100.0%; Pred. No. 1.2e-257;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
 |||||
 Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy 61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIYRCLVG 120
 |||||
 Db 61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIYRCLVG 120

Qy 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 |||||
 Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAEESDNVDSADAEEEDSDVWGGADTDYADGSEDKVVEVAEEEEVAEVEEE 240
 |||||
 Db 181 GVEFVCCPLAEESDNVDSADAEEEDSDVWGGADTDYADGSEDKVVEVAEEEEVAEVEEE 240

Qy 241 EADDDDEDDGDEVEEEAEPEYEEATERTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300
 |||||
 Db 241 EADDDDEDDGDEVEEEAEPEYEEATERTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300

Qy 301 RAMISRWFVDVTEGKCAPFFYGGCGGNRNPNFDTEEYCMVCGSAMSQSLKTTQEPLARD 360
 |||||
 Db 301 RAMISRWFVDVTEGKCAPFFYGGCGGNRNPNFDTEEYCMVCGSAMSQSLKTTQEPLARD 360

Qy 361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420

Db	361		PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Qy	421		KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	421		KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Qy	481		QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481		QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541		MNQSLSLLYNVPVAEELQDEVDLLQKEQNYSDDLANMISEPRISYNDALMPSLTET	600
Db	541		MNQSLSLLYNVPVAEELQDEVDLLQKEQNYSDDLANMISEPRISYNDALMPSLTET	600
Qy	601		KTTVELLPVNGEFSDDLQPWHSFGADSVPAANTENEVEPVDPARPAADRGLTTRPGSGLTN	660
Db	601		KTTVELLPVNGEFSDDLQPWHSFGADSVPAANTENEVEPVDPARPAADRGLTTRPGSGLTN	660
Qy	661		IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661		IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721		VMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770
Db	721		VMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770

RESULT 11

US-09-869-414-55

; Sequence 55, Application US/09869414

; Publication No. US20030077226A1

; GENERAL INFORMATION:

; APPLICANT: Beinkowski et al.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: 28341/6280M

; CURRENT APPLICATION NUMBER: US/09/869,414

; CURRENT FILING DATE: 2001-06-27

; PRIOR APPLICATION NUMBER: 09/416,901

; PRIOR FILING DATE: 1999-10-13

; PRIOR APPLICATION NUMBER: 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 55

; LENGTH: 770

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-869-414-55

US-09-548-366-55

; Publication No. US20030104365A1

; APPLICANT: Gurney, Mark E.

APPLICANT: Bienkowski, Michael J.

; APPLICANT: Heinrikson, Robert L.

; APPLICANT: Parodi, Luis A.

; APPLICANT: Yan, Riqiang

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR,
AND

; TITLE OF INVENTION: USES THEREFOR

; FILE REFERENCE: 28341/6280A

CURRENT APPLICATION NUMBER: US/09/548,366

CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 09/404,133

; PRIOR FILING DATE: 1999-09-23

PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

PRIOR APPLICATION NUMBER: 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 55

; LENGTH: 770

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-548-366-55

Query Match 100.0%; Score 4058; DB 11; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.2e-257;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRINMHMNVQNGKWDSDPSGK 60

Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy 61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIYRCLVG 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

Qy 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 121 EFVSDALLVPDKCKFLHOERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAEESDNVDSADAEEDDSDVWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 181 GVEFVCCPLAEESDNVDSADAEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE 240

Qy 241 EADDDDEDEDGDEVEEEAEPEYEEATERTTSIATTTTTTTSVEEVVREVCSEQAETGPC 300
 |
 |
 |

Db 241 EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300


```

;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/10/335,035
;      FILING DATE: 30-Dec-2002
;      CLASSIFICATION: 435
;
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US/08/665,649
;      FILING DATE: 18-JUN-1996
;      APPLICATION NUMBER: US 08/437,067
;      FILING DATE: 09-MAY-1995
;
;      ATTORNEY/AGENT INFORMATION:
;      NAME: STORELLA ESQ., JOHN R.
;      REGISTRATION NUMBER: 32,944
;      REFERENCE/DOCKET NUMBER: 15270-000650
;
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (415) 326-2400
;      TELEFAX: (415) 576-0300
;
;      INFORMATION FOR SEQ ID NO: 3:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 770 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
;      SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-335-035-3

```

```

Query Match          100.0%;  Score 4058;  DB 12;  Length 770;
Best Local Similarity 100.0%;  Pred. No. 1.2e-257;
Matches 770;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      1  MLPGLIALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MLPGLIALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

Qy    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181  GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241  EADDDDEDDGDEVEEEAEEPYEEATERTTTSIATTTTTTTTSESVEEVVREVCSEQAETGPC 300
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  EADDDDEDDGDEVEEEAEEPYEEATERTTTSIATTTTTTTTSESVEEVVREVCSEQAETGPC 300

Qy    301  RAMISRWFYFDVTEGKCAPFFYGGCGGNRNRFDTTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  RAMISRWFYFDVTEGKCAPFFYGGCGGNRNRFDTTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361  PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420

```

Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNDAIMPSTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNDAIMPSTET	600
Qy	601	KTTVELLPVNGEFSLLDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLLDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFPAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFPAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMN	770
Db	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMN	770

US-10-223-809A-2

; Publication No. US20030157117A1

APPLICANT: Pharmexa A/S

; TITLE OF INVENTION: No. US20030157117A1e1 Method for Down-Regulation of Amyloid

; CURRENT APPLICATION NUMBER: US/10/223,809A

; PRIOR APPLICATION NUMBER: US 60/337,543

PRIOR APPLICATION NUMBER: US 60/373,027

; PRIOR APPLICATION NUMBER: DE 2001 01231

; PRIOR APPLICATION NUMBER: DE 2002 0058

; NUMBER OF SEO ID NOS: 17

; SEQ ID NO 2

; TYPE: PRT

; FEATURE:

; LOCATION: (2098) .. (2169)

FEATURE:

```

; NAME/KEY: misc_feature
; LOCATION: (2014)..(2313)
; OTHER INFORMATION: nucleotides encoding C-100
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2016)..(2144)
; OTHER INFORMATION: Abeta 42/43
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2014)..(2142)
; OTHER INFORMATION: Abeta 42/43
US-10-223-809A-2

```

```

Query Match          100.0%; Score 4058; DB 12; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.2e-257;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIRCLVG 120
      |||
Db     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIRCLVG 120

Qy    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241 EADDDDEDDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241 EADDDDEDDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300

Qy    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420
      |||
Db    361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420

Qy    421 KNLPKADKKAVIQHFQEKVESLEQEAAENERQQLVETHMARVEAMLNDRRRLALENYITAL 480
      |||
Db    421 KNLPKADKKAVIQHFQEKVESLEQEAAENERQQLVETHMARVEAMLNDRRRLALENYITAL 480

Qy    481 QAVPPRRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
      |||
Db    481 QAVPPRRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Qy    541 MNQSLSLLYNVPAAVEEIQDEVEDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET 600
      |||
Db    541 MNQSLSLLYNVPAAVEEIQDEVEDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET 600

```

Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV	PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV	PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEI SEVKMDAEFRHDSGYEVHHQKL	VFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEI SEVKMDAEFRHDSGYEVHHQKL	VFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLS	KMQQNGYENPTYKFFEQQMN	770
Db	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLS	KMQQNGYENPTYKFFEQQMN	770

RESULT 15

US-10-010-942B-38

```

; Sequence 38, Application US/10010942B
; Publication No. US20030165496A1
; GENERAL INFORMATION:
; APPLICANT: Basi, Guriq
; APPLICANT: Saldanha, Jose
; APPLICANT: Yednock, Ted
; TITLE OF INVENTION: HUMANIZED ANTIBODIES THAT RECOGNIZE
; TITLE OF INVENTION: BETA AMYLOID PEPTIDE
; FILE REFERENCE: ELN-002
; CURRENT APPLICATION NUMBER: US/10/010,942B
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/251,892
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-10-010-942B-38

```

Query Match          100.0%; Score 4058; DB 12; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.2e-257;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGR	LNMHMNVQNGKWDSDPSGTK	60
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGR	LNMHMNVQNGKWDSDPSGTK	60
Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI	PYRCLVG	120
Db	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI	PYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180	
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180	
Qy	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240	
Db	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240	
Qy	241	EADDDDEDDEGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC	300	

Db	241	EADDDDEDEDGDEVEEEEAEPEYEEATERTTTSIATTTTTTTESVVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRISYGNDALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRISYGNDALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMN	770
Db	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMN	770

Search completed: December 13, 2003, 13:36:48
 Job time : 89.4359 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 47.0085 Seconds
(without alignments)
693.052 Million cell updates/sec

Title: US-09-785-215-2
Perfect score: 4058
Sequence: 1 MLPGLALLLLLAAWTARALEV.....KMQQNGYENPTYKFFEQMQN 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4058	100.0	770	1 US-08-133-248-8	Sequence 8, Appli
2	4058	100.0	770	1 US-08-231-940-1	Sequence 1, Appli
3	4058	100.0	770	2 US-08-641-774-1	Sequence 1, Appli
4	4058	100.0	770	2 US-08-104-165-3	Sequence 3, Appli
5	4058	100.0	770	3 US-08-464-250-3	Sequence 3, Appli
6	4058	100.0	770	4 US-08-464-250-3	Sequence 3, Appli
7	4058	100.0	770	4 US-09-548-372D-55	Sequence 55, Appl
8	4058	100.0	770	4 US-09-548-367D-55	Sequence 55, Appl
9	4058	100.0	770	4 US-09-551-853D-55	Sequence 55, Appl
10	4058	100.0	772	4 US-09-548-372D-59	Sequence 59, Appl
11	4058	100.0	772	4 US-09-548-367D-59	Sequence 59, Appl

12	4058	100.0	772	4	US-09-551-853D-59	Sequence 59, Appl
13	3943.5	97.2	751	1	US-08-123-702-4	Sequence 4, Appli
14	3943.5	97.2	751	2	US-08-104-165-2	Sequence 2, Appli
15	3943.5	97.2	751	2	US-08-422-333-2	Sequence 21, Appli
16	3943.5	97.2	751	2	US-08-422-333-21	Sequence 2, Appli
17	3943.5	97.2	751	3	US-08-464-250-2	Sequence 2, Appli
18	3943.5	97.2	751	4	US-08-464-250-2	Sequence 2, Appli
19	3943.5	97.2	751	4	US-08-832-867-5	Sequence 5, Appli
20	3943.5	97.2	751	4	US-09-548-372D-57	Sequence 57, Appl
21	3943.5	97.2	751	4	US-09-548-367D-57	Sequence 57, Appl
22	3943.5	97.2	751	4	US-09-551-853D-57	Sequence 57, Appl
23	3943.5	97.2	751	6	5187153-2	Patent No. 5187153
24	3943.5	97.2	751	6	5223482-2	Patent No. 5223482
25	3943.5	97.2	753	4	US-09-548-372D-61	Sequence 61, Appl
26	3943.5	97.2	753	4	US-09-548-367D-61	Sequence 61, Appl
27	3943.5	97.2	753	4	US-09-551-853D-61	Sequence 61, Appl
28	3937.5	97.0	751	6	5220013-2	Patent No. 5220013
29	3590.5	88.5	695	1	US-08-123-702-2	Sequence 2, Appli
30	3590.5	88.5	695	2	US-08-104-165-1	Sequence 1, Appli
31	3590.5	88.5	695	3	US-08-464-250-1	Sequence 1, Appli
32	3590.5	88.5	695	4	US-08-464-250-1	Sequence 1, Appli
33	3590.5	88.5	695	4	US-09-458-481B-7	Sequence 7, Appli
34	3590.5	88.5	695	4	US-09-458-481B-8	Sequence 8, Appli
35	3590.5	88.5	695	4	US-09-548-372D-10	Sequence 10, Appl
36	3590.5	88.5	695	4	US-09-548-367D-10	Sequence 10, Appl
37	3590.5	88.5	695	4	US-09-551-853D-10	Sequence 10, Appl
38	3590.5	88.5	695	6	5218100-2	Patent No. 5218100
39	3590.5	88.5	697	4	US-09-548-372D-16	Sequence 16, Appl
40	3590.5	88.5	697	4	US-09-548-367D-16	Sequence 16, Appl
41	3590.5	88.5	697	4	US-09-551-853D-16	Sequence 16, Appl
42	3585.5	88.4	695	4	US-09-548-372D-14	Sequence 14, Appl
43	3585.5	88.4	695	4	US-09-548-367D-14	Sequence 14, Appl
44	3585.5	88.4	695	4	US-09-551-853D-14	Sequence 14, Appl
45	3585.5	88.4	697	4	US-09-548-372D-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-08-133-248-8

; Sequence 8, Application US/08133248

; Patent No. 5525714

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: MUTATED FORM OF THE BETA-AMYLOID PRECURSOR

; TITLE OF INVENTION: PROTEIN GENE

; NUMBER OF SEQUENCES: 8

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/133,248

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

```

;      LENGTH:  770 amino acids
;      TYPE:    amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE:  protein
US-08-133-248-8

```

Query Match 100.0%; Score 4058; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 2.3e-273;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIRCLVG	120
Db	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE	240
Qy	241	EADDDDEDEDGDEVEEEAEFPYEEATERTTTSIATTTTTTTTESVEEVVREVCSEAETGPC	300
Db	241	EADDDDEDEDGDEVEEEAEFPYEEATERTTTSIATTTTTTTTESVEEVVREVCSEAETGPC	300
Qy	301	RAMISRWFYDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFYDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNDALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNDALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPV DARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPV DARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720

Qy 721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEOMQN 770
 Db 721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEOMQN 770

RESULT 2

US-08-231-940-1

; Sequence 1, Application US/08231940
 ; Patent No. 5550216
 ; GENERAL INFORMATION:
 ; APPLICANT: MIYAZAKI, Kaoru
 ; TITLE OF INVENTION: GELATINASE A INHIBITOR
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BROWDY AND NEIMARK
 ; STREET: 419 Seventh Street, N.W., Suite 300
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/231,940
 ; FILING DATE: 25-APR-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 5-120457
 ; FILING DATE: 26-APR-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-62129
 ; FILING DATE: 08-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: NEIMARK, Sheridan
 ; REGISTRATION NUMBER: 20,520
 ; REFERENCE/DOCKET NUMBER: MIYAZAKI=4
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-628-5197
 ; TELEFAX: 202-737-3528
 ; TELEX: 248633
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 770 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide

US-08-231-940-1

Query Match 100.0%; Score 4058; DB 1; Length 770;
 Best Local Similarity 100.0%; Pred. No. 2.3e-273;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGR LNMH MNVQNGKWDS DPSG TK	60
Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG	120
Db	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGM L L P C G I D K F R	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGM L L P C G I D K F R	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Qy	241	EADDDDEDEDGDEVEEEEAEPYEEATER TTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Db	241	EADDDDEDEDGDEVEEEEAEPYEEATER TTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVAEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVAEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAI IGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAI IGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFE QM QN	770
Db	721	VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFE QM QN	770

RESULT 3

US-08-641-774-1

; Sequence 1, Application US/08641774

; Patent No. 5843695

; GENERAL INFORMATION:

Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
 |||

Db 181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
 |||

QY 241 EADDDDEDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300
 |||

Db 241 EADDDDEDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300
 |||

QY 301 RAMISRWFYFDVTEGKCAPFFYGGCGNRNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
 |||

Db 301 RAMISRWFYFDVTEGKCAPFFYGGCGNRNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
 |||

QY 361 PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420
 |||

Db 361 PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420
 |||

QY 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480
 |||

Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480
 |||

QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 |||

Db 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 |||

QY 541 MNQSLSLLYNVPVAVAEEIQDEVDLLQKEQNYSDDLANMISEPRISYGNDALMPSLTET 600
 |||

Db 541 MNQSLSLLYNVPVAVAEEIQDEVDLLQKEQNYSDDLANMISEPRISYGNDALMPSLTET 600
 |||

QY 601 KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 |||

Db 601 KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 |||

QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 |||

Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 |||

QY 721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 |||

Db 721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 |||

RESULT 4

US-08-104-165-3

; Sequence 3, Application US/08104165

; Patent No. 5877015

; GENERAL INFORMATION:

; APPLICANT: HARDY, John Anthony

; APPLICANT: GOATE, Alison Mary

; APPLICANT: MULLAN, Michael John

; APPLICANT: CHARTIER-HARLIN, Marie-Christine

; APPLICANT: OWEN, Michael John

; TITLE OF INVENTION: Test and Model for Alzheimer's Disease

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Khourie and Crew

```

; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/104,165
; FILING DATE: 21-JAN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-104-165-3

```

```

Query Match          100.0%; Score 4058; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 2.3e-273;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDPSGTK 60
      |||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDPSGTK 60

Qy     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIRCLVG 120
      |||
Db     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIRCLVG 120

Qy    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241 EADDDDEDDEDGDEVEEEEAEPEYEEATERTTTSIATTTTTTTTSESVEEVVREVCSEQAETGPC 300
      |||
Db    241 EADDDDEDDEDGDEVEEEEAEPEYEEATERTTTSIATTTTTTTTSESVEEVVREVCSEQAETGPC 300

```


Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770
Db	721	VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770

RESULT 5

US-08-464-250-3

; Sequence 3, Application US/08464250

; Patent No. 6107542

; GENERAL INFORMATION:

; APPLICANT: HARDY, John Anthony

; APPLICANT: GOATE, Alison Mary

; APPLICANT: MULLAN, Michael John

; APPLICANT: CHARTIER-HARLIN, Marie-Christine

; APPLICANT: OWEN, Michael John

; TITLE OF INVENTION: Test and Model for Alzheimer's Disease

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Khourie and Crew

; STREET: 379 Lytton Avenue

; CITY: Palo Alto

; STATE: California

; COUNTRY: US

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,250
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/104,165
; FILING DATE: 21-JAN-1992
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-464-250-3

```

```

Query Match          100.0%; Score 4058; DB 3; Length 770;
Best Local Similarity 100.0%; Pred. No. 2.3e-273;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
      |||
Db     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

Qy    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181 GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181 GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241 EADDDDEDDEDGDEVEEEAEPEYEEATERTTISIATTTTTTTTSESVEEVVREVCSEQAETGPC 300
      |||
Db    241 EADDDDEDDEDGDEVEEEAEPEYEEATERTTISIATTTTTTTTSESVEEVVREVCSEQAETGPC 300

Qy    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
      |||

```

Db 361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
 QY 421 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 Db 421 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 Db 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 QY 541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNLALMPSLTET 600
 Db 541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLNMISEPRISYGNLALMPSLTET 600
 QY 601 KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDPARPAADRGLTTRPGSGLTN 660
 Db 601 KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDPARPAADRGLTTRPGSGLTN 660
 QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 QY 721 VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 Db 721 VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 6

US-08-464-250-3

; Sequence 3, Application US/08464250

; Patent No. 6300540

; GENERAL INFORMATION:

; APPLICANT: HARDY, John Anthony

; GOATE, Alison Mary

; MULLAN, Michael John

; CHARTIER-HARLIN, Marie-Christine

; OWEN, Michael John

; TITLE OF INVENTION: Test and Model for Alzheimer's Disease

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Khourie and Crew

; STREET: 379 Lytton Avenue

; CITY: Palo Alto

; STATE: California

; COUNTRY: US

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/464,250

; FILING DATE: 05-Jun-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/104,165

; FILING DATE: 21-JAN-1992

```

; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-464-250-3

```

```

Query Match          100.0%; Score 4058; DB 4; Length 770;
Best Local Similarity 100.0%; Pred. No. 2.3e-273;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
      |||
Db     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

Qy    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMILLPCGIDKFR 180
      |||
Db    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMILLPCGIDKFR 180

Qy    181  GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181  GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241  EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241  EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300

Qy    301  RAMISRWFYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301  RAMISRWFYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361  PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
      |||
Db    361  PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420

Qy    421  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480
      |||
Db    421  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480

```

```

Qy      481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
      |||
Db      481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Qy      541 MNQSLSLLYNVPVAEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET 600
      |||
Db      541 MNQSLSLLYNVPVAEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET 600

Qy      601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
      |||
Db      601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Qy      661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
      |||
Db      661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720

Qy      721 VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
      |||
Db      721 VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

```

RESULT 7

US-09-548-372D-55

; Sequence 55, Application US/09548372D

; Patent No. 6420534

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/6280I

; CURRENT APPLICATION NUMBER: US/09/548,372D

; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 55

; LENGTH: 770

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-548-372D-55

Query Match 100.0%; Score 4058; DB 4; Length 770;

Best Local Similarity 100.0%; Pred. No. 2.3e-273;

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

```

Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG	120
Db	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEDDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEDDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE	240
Qy	241	EADDDDEDEDGDEVEEEAEOPYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Db	241	EADDDDEDEDGDEVEEEAEOPYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDPELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDPELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKL VFFAEDVGSNKGAI IGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKL VFFAEDVGSNKGAI IGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQRN	770
Db	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQRN	770

RESULT 8

US-09-548-367D-55

; Sequence 55, Application US/09548367D

; Patent No. 6440698

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

```

; TITLE OF INVENTION:  THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-55

```

```

Query Match          100.0%; Score 4058; DB 4; Length 770;
Best Local Similarity 100.0%; Pred. No. 2.3e-273;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDPSGTK 60
      |||
Db      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDPSGTK 60

Qy     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
      |||
Db     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

Qy    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181  GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181  GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241  EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTTSVEEVVREVCSEQAETGPC 300
      |||
Db    241  EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTTSVEEVVREVCSEQAETGPC 300

Qy    301  RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301  RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361  PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
      |||
Db    361  PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420

Qy    421  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
      |||
Db    421  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480

Qy    481  QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

```

```

Db      481  |||||QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
QY      541  MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGN DALMPSLTET 600
Db      541  |||||MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGN DALMPSLTET 600
QY      601  KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db      601  |||||KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
QY      661  IKTEEISEVKMDAEFRHDSGYEVHHQKL VFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
Db      661  |||||IKTEEISEVKMDAEFRHDSGYEVHHQKL VFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
QY      721  VMLKKKQYTSIHGVEVDAAVTPEERHLSKM QQNGYENPTYKFFEQMQN 770
Db      721  |||||VMLKKKQYTSIHGVEVDAAVTPEERHLSKM QQNGYENPTYKFFEQMQN 770

```

RESULT 9

US-09-551-853D-55

; Sequence 55, Application US/09551853D

; Patent No. 6500667

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/6280L

; CURRENT APPLICATION NUMBER: US/09/551,853D

; CURRENT FILING DATE: 2000-04-18

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 55

; LENGTH: 770

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-551-853D-55

Query Match 100.0%; Score 4058; DB 4; Length 770;

Best Local Similarity 100.0%; Pred. No. 2.3e-273;

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMH MNVQNGKWDSDPSGTK 60
Db      1  |||||MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMH MNVQNGKWDSDPSGTK 60
QY      61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

```


Db	61	TCIDTKEGILQYCQEVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEDDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEDDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE	240
Qy	241	EADDDDEDEDGDEVEEEAEAPYEEATERTTSIATTTTTTTTSEVVEVVREVCSEQAETGPC	300
Db	241	EADDDDEDEDGDEVEEEAEAPYEEATERTTSIATTTTTTTTSEVVEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDLLQKEQNYSDDLANMISEPRISYGNLALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDLLQKEQNYSDDLANMISEPRISYGNLALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVPFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVPFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMN	770
Db	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMN	770

RESULT 10

US-09-548-372D-59

; Sequence 59, Application US/09548372D

; Patent No. 6420534

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-59

Query Match 100.0%; Score 4058; DB 4; Length 772;
Best Local Similarity 100.0%; Pred. No. 2.3e-273;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
| | | | |
Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy 61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
| | | | |
Db 61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

Qy 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
| | | | |
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAEESDNVDSADAEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
| | | | |
Db 181 GVEFVCCPLAEESDNVDSADAEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy 241 EADDDDEDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTSESVEEVVREVCSEQAETGPC 300
| | | | |
Db 241 EADDDDEDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTSESVEEVVREVCSEQAETGPC 300

Qy 301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
| | | | |
Db 301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy 361 PVKLP'TTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
| | | | |
Db 361 PVKLP'TTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420

Qy 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480
| | | | |
Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480

Qy 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
| | | | |

Db 61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
 Db 181 GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
 QY 241 EADDDDEDGEDGDEVEEEAEAPYEEATERTTTSIATTTTTTTTSESVEEVVREVCSEQAETGPC 300
 Db 241 EADDDDEDGEDGDEVEEEAEAPYEEATERTTTSIATTTTTTTTSESVEEVVREVCSEQAETGPC 300
 QY 301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
 Db 301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
 QY 361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
 Db 361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
 QY 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 Db 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 QY 541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDAIMPSTET 600
 Db 541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDAIMPSTET 600
 QY 601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 Db 601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 QY 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 12

US-09-551-853D-59

; Sequence 59, Application US/09551853D

; Patent No. 6500667

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/6280L

Qy 541 MNQSLSLLYNVPVAEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 541 MNQSLSLLYNVPVAEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET 600
 Qy 601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
 Qy 721 VMLKKKQYTSIHG VVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 721 VMLKKKQYTSIHG VVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 13

US-08-123-702-4

; Sequence 4, Application US/08123702

; Patent No. 5604131

; GENERAL INFORMATION:

; APPLICANT: Wadsworth, Samuel

; APPLICANT: Snyder, Benjamin

; APPLICANT: Reddy, Vermuri, B.

; APPLICANT: Wei, Chamer

; TITLE OF INVENTION: A cDNA Genomic Hybrid Sequence Encoding APP770

; Patent No. 5604131

; TITLE OF INVENTION: Containing a Genomic DNA Insert of the KI and OX-2 Regions

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Patrea L. Pabst

; STREET: 2800 One Atlantic Center

; STREET: 1201 West Peachtree Street

; CITY: Atlanta

; STATE: GA

; COUNTRY: USA

; ZIP: 30309-3450

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/123,702

; FILING DATE: 17-SEPT-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Pabst, Patrea L.

; REGISTRATION NUMBER: 31,284

; REFERENCE/DOCKET NUMBER: TS1121

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (404)-873-8794

; TELEFAX: (404)-873-8795

; INFORMATION FOR SEQ ID NO: 4:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-123-702-4

```

```

Query Match          97.2%; Score 3943.5; DB 1; Length 751;
Best Local Similarity 97.4%; Pred. No. 1.9e-265;
Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

```

```

Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDPSGTK 60
      |||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDPSGTK 60

Qy     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
      |||
Db     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

Qy    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181 GVEFVCCPLAEESDNVDSADAEEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE 240
      |||
Db    181 GVEFVCCPLAEESDNVDSADAEEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE 240

Qy    241 EADDDDEDDGDEVEEEAEPEYEEATERTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241 EADDDDEDDGDEVEEEAEPEYEEATERTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300

Qy    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTTEEYCMAVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTTEEYCMAVCGSA----- 344

Qy    361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEEAERQA 420
      :|||
Db    345 ---IPTTAASTPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEEAERQA 401

Qy    421 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480
      |||
Db    402 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL 461

Qy    481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
      |||
Db    462 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 521

Qy    541 MNQSLSLLYNVPAAVAEEIQDEVDLLQKEQNYSDVLANMISEPRISYGNDAIMPSTLTET 600
      |||
Db    522 MNQSLSLLYNVPAAVAEEIQDEVDLLQKEQNYSDVLANMISEPRISYGNDAIMPSTLTET 581

Qy    601 KTTVELLPVNGEFSDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 660
      |||
Db    582 KTTVELLPVNGEFSDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 641

Qy    661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
      |||

```

Db 642 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 701

QY 721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 |||

Db 702 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 751

RESULT 14

US-08-104-165-2

; Sequence 2, Application US/08104165

; Patent No. 5877015

; GENERAL INFORMATION:

; APPLICANT: HARDY, John Anthony

; APPLICANT: GOATE, Alison Mary

; APPLICANT: MULLAN, Michael John

; APPLICANT: CHARTIER-HARLIN, Marie-Christine

; APPLICANT: OWEN, Michael John

; TITLE OF INVENTION: Test and Model for Alzheimer's Disease

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Kourie and Crew

; STREET: 379 Lytton Avenue

; CITY: Palo Alto

; STATE: California

; COUNTRY: US

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/104,165

; FILING DATE: 21-JAN-1992

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 9101307.8

; FILING DATE: 21-JAN-1991

; APPLICATION NUMBER: 9118445.7

; FILING DATE: 28-AUG-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Liebeschuetz, Joe

; REGISTRATION NUMBER: 37,505

; REFERENCE/DOCKET NUMBER: 16163-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 326-2400

; TELEFAX: (415) 326-2422

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 751 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-104-165-2

Query Match 97.2%; Score 3943.5; DB 2; Length 751;
 Best Local Similarity 97.4%; Pred. No. 1.9e-265;

Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

```

QY      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDSPSGTK 60
      |||
Db      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWSDSPSGTK 60

QY     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
      |||
Db     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

QY    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

QY    181  GVEFVCCPLAEESDNVDSADAEEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181  GVEFVCCPLAEESDNVDSADAEEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

QY    241  EADDDDEDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241  EADDDDEDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300

QY    301  RAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301  RAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSA----- 344

QY    361  PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420
      :|||
Db    345  ---IPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 401

QY    421  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480
      |||
Db    402  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 461

QY    481  QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
      |||
Db    462  QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 521

QY    541  MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET 600
      |||
Db    522  MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET 581

QY    601  KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 660
      |||
Db    582  KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 641

QY    661  IKTEEISEVKMDAEFRHDSGYEVVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
      |||
Db    642  IKTEEISEVKMDAEFRHDSGYEVVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 701

QY    721  VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
      |||
Db    702  VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 751

```

```

; Sequence 2, Application US/08422333
; Patent No. 5912410
; GENERAL INFORMATION:
;   APPLICANT: CORDELL, Barbara L.
;   TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMAL DISPLAYING
;   TITLE OF INVENTION: THE AMYLOID-FORMING PATHOLOGY OF ALZHEIMER'S DISEASE
;   NUMBER OF SEQUENCES: 30
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Scios, Inc.
;     STREET: 2450 Bayshore Parkway
;     CITY: Mountain View
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94043
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/422,333
;     FILING DATE: 13-APR-1995
;     CLASSIFICATION: 800
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Shearer, Peter R.
;     REGISTRATION NUMBER: 28,117
;     REFERENCE/DOCKET NUMBER: 21900-28048.00
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (415) 966-1550
;     TELEFAX: (415) 968-2438
;   INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 751 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-422-333-2

```

```

Query Match          97.2%;  Score 3943.5;  DB 2;  Length 751;
Best Local Similarity 97.4%;  Pred. No. 1.9e-265;
Matches 750;  Conservative 1;  Mismatches 0;  Indels 19;  Gaps 1;

```

```

Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGK 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGK 60

Qy     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

Qy    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181 GVEFVCCPLAEESDNVDSADAEEDDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GVEFVCCPLAEESDNVDSADAEEDDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

```


OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 127.863 Seconds
(without alignments)
955.861 Million cell updates/sec

Title: US-09-785-215-2
Perfect score: 4058
Sequence: 1 MLPGLALLLLAAWTARALEV.....KMQQNGYENPTYKFFEQMQN 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	4058	100.0	770	10	AAP94775		Novel amyloid prec
2	4058	100.0	770	14	AAR41546		Mutated APP770 exo
3	4058	100.0	770	15	AAR63442		Amyloid protein pr
4	4058	100.0	770	19	AAW40130		Human APP770 prote
5	4058	100.0	770	20	AAW97996		Human amyloid prec
6	4058	100.0	770	22	AAE10648		Human amyloid prot
7	4058	100.0	770	22	AAE11762		Human amyloid prec
8	4058	100.0	770	22	AAE06893		Human amyloid prec
9	4058	100.0	770	22	AAU06622		Human partial Amyl
10	4058	100.0	770	22	AAE02600		Human amyloid prec
11	4058	100.0	770	23	ABG94279		Amyloid beta prote
12	4058	100.0	770	23	ABG32723		Human amyloid prec
13	4058	100.0	770	23	ABG80591		Human amyloid beta
14	4058	100.0	770	23	ABG76936		Humanised antibody
15	4058	100.0	770	23	ABB78609		Human APP770 prote
16	4058	100.0	770	23	AAG68317		Human amyloid prec
17	4058	100.0	770	24	ABP72693		Human amyloid prec
18	4058	100.0	770	24	ABP97885		Amino acid sequenc
19	4058	100.0	772	22	AAE10650		Human amyloid prot
20	4058	100.0	772	22	AAE06895		Human amyloid prec
21	4058	100.0	772	22	AAU06624		Human Amyloid prec
22	4058	100.0	772	22	AAU07223		Human beta-amyloid
23	4058	100.0	772	22	AAE02602		Human amyloid prec
24	4058	100.0	772	23	ABB78611		Human APP770-KK pr
25	4054	99.9	770	15	AAR62505		Amyloid precursor
26	4053	99.9	770	18	AAW19500		APP770 mutant A-be
27	4053	99.9	770	18	AAW19497		APP770 mutant A-be
28	4053	99.9	770	18	AAW19485		APP770 mutant A-be
29	4053	99.9	770	18	AAW19482		APP770 mutant A-be
30	4053	99.9	770	22	AAE06913		Human amyloid prec
31	4053	99.9	770	23	ABB78008		Amino acid sequenc
32	4050	99.8	770	22	AAE06912		Human amyloid prec
33	4047	99.7	768	23	AAU80959		Human amyloid beta
34	4047	99.7	770	18	AAW19506		APP770 mutant A-be
35	4047	99.7	770	18	AAW19491		APP770 mutant A-be
36	4046	99.7	770	13	AAR26340		APP770. Homo sapi
37	4046	99.7	770	18	AAW19488		APP770 mutant A-be
38	4046	99.7	770	18	AAW19503		APP770 mutant A-be
39	4039	99.5	770	11	AAR05717		NAP gene product a
40	3943.5	97.2	751	12	AAR10022		Beta-amyloid-relat
41	3943.5	97.2	751	13	AAR20328		Sequence encoded b
42	3943.5	97.2	751	20	AAV08615		Human beta-amyloid
43	3943.5	97.2	751	20	AAV08605		Human beta-amyloid
44	3943.5	97.2	751	22	AAE10649		Human amyloid prot
45	3943.5	97.2	751	22	AAE06894		Human amyloid prec

ALIGNMENTS

RESULT 1

AAP94775

ID AAP94775 standard; protein; 770 AA.

XX

AC AAP94775;

XX

DT 25-MAR-2003 (updated)

DT 05-JUL-1990 (first entry)

XX

DE Novel amyloid precursor protein (NAP).

XX

KW SPAP; amyloid precursor protein; dysbolism; INS76; ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Active-site 289..345

FT /*tag= a

FT /*label=INS76

XX

PN EP304013-A.

XX

PD 22-FEB-1989.

XX

PF 16-AUG-1988; 88EP-0113283.

XX

PR 15-AUG-1987; 87JP-0203298.

PR 21-AUG-1987; 87JP-0207995.

PR 18-NOV-1987; 87JP-0291404.

PR 11-DEC-1987; 87JP-0313228.

PR 05-FEB-1988; 88JP-0025260.

PR 10-FEB-1988; 88JP-0029366.

PR 19-FEB-1988; 88JP-0037905.

PR 25-MAY-1988; 88JP-0125660.

XX

PA (ASAH) ASahi KASEI KOGYO KK.

XX

PI Kitaguchi N, Takahashi Y, Tokushima Y, Itoh H;

XX

DR WPI; 1989-055458/08.

DR N-PSDB; AAN91049.

XX

PT Human senile plaque amyloid precursor protein and DNA -

PT used for study and diagnosis of dysbolism in the central nervous

PT system.

XX

PS Disclosure; Page -; 108pp; English.

XX

CC AA. sequence from 289 to 345 encodes INS76 which is claimed as a senile
 CC plaque amyloid precursor (SPAP), useful for diagnosis of dysbolism in the
 CC CNS such as senile dementia.
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX

SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 10; Length 770;
 Best Local Similarity 100.0%; Pred. No. 9.2e-289;

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
      |||
Db     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

Qy    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181 GVEFVCCPLAEESDNVDSADAEEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181 GVEFVCCPLAEESDNVDSADAEEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241 EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241 EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300

Qy    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361 PVKLPPTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420
      |||
Db    361 PVKLPPTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA 420

Qy    421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480
      |||
Db    421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480

Qy    481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAQIRSQVMTHLRVIYER 540
      |||
Db    481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAQIRSQVMTHLRVIYER 540

Qy    541 MNQSLSLLYNVPAAVEEIQDEVDELLOKEQNYSDDLANMISEPRISYGNDALMPSLTET 600
      |||
Db    541 MNQSLSLLYNVPAAVEEIQDEVDELLOKEQNYSDDLANMISEPRISYGNDALMPSLTET 600

Qy    601 KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 660
      |||
Db    601 KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Qy    661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
      |||
Db    661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720

Qy    721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMN 770
      |||
Db    721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMN 770

```

RESULT 2
AAR41546

ID AAR41546 standard; Protein; 770 AA.
XX
AC AAR41546;
XX
DT 25-MAR-2003 (updated)
DT 15-MAR-1994 (first entry)
XX
DE Mutated APP770 exon 17 protein fragment.
XX
KW Probe; mutation; exon 17; amyloid precursor protein; APP770;
KW substitution; progressive presenile dementia; Alzheimer's disease;
KW cerebral haemorrhage; cerebral amyloid angiopathy; 4 kD fragment;
KW blood vessels; brain parenchyma; assay; processing; plaque.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 692
FT /note= "Position of mutation which causes abnormal
FT processing of APP770"
XX
PN EP561087-A1.
XX
PD 22-SEP-1993.
XX
PF 20-MAR-1992; 92EP-0400771.
XX
PR 20-MAR-1992; 92EP-0400771.
XX
PA (INNO-) INNOGENETICS NV SA.
XX
PI Cras P, Hendriks L, Martin J, Van Broeckhoven C;
XX
DR WPI; 1993-296442/38.
DR N-PSDB; AAQ48860.
XX
PT New mutant form of beta-amyloid polypeptide - related to
PT development of cerebral haemorrhage and Alzheimer's disease, also
PT corresp. nucleic acid, vectors, host cells and antibodies
XX
PS Disclosure; Fig 4; 2lpp; English.
XX
CC This sequence is encoded by exon 17 of the amyloid precursor protein
CC APP770. A mutation at position 2075 of the DNA sequence, a C>G
CC substitution, causes the substitution of Ala for Gly at codon 692.
CC This mutation has been detected in related patients with progressive
CC presenile dementia (Alzheimer's disease) or cerebral haemorrhage due
CC to cerebral amyloid angiopathy. It may be responsible for the
CC deposition of a 4 kD proteolytic fragment of APP in blood vessel
CC walls and brain parenchyma. Probes specific for the mutation (see
CC also AAQ48858-59) can be used to assay mRNA encoding substances which
CC cause abnormal processing of APP related to plaque formation, and to
CC detect this specific mutation.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 14; Length 770;
Best Local Similarity 100.0%; Pred. No. 9.2e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
      |||
Db     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

Qy    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181 GVEFVCCPLAEESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181 GVEFVCCPLAEESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241 EADDDDEDDGDEVEEEAEPEYEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241 EADDDDEDDGDEVEEEAEPEYEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300

Qy    301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420
      |||
Db    361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420

Qy    421 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQVLVETHMARVEAMLNDRRRLALENYITAL 480
      |||
Db    421 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQVLVETHMARVEAMLNDRRRLALENYITAL 480

Qy    481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
      |||
Db    481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Qy    541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRI SYGNDALMPSLTET 600
      |||
Db    541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRI SYGNDALMPSLTET 600

Qy    601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
      |||
Db    601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Qy    661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAI IGLMVGGVVIATVIVITL 720
      |||
Db    661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAI IGLMVGGVVIATVIVITL 720

Qy    721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
      |||
Db    721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
```

RESULT 3

AAR63442

ID AAR63442 standard; protein; 770 AA.

XX

AC AAR63442;

XX

DT 25-MAR-2003 (updated)

DT 30-JUN-1995 (first entry)

XX

DE Amyloid protein precursor APP.

XX

KW Amyloid protein precursor; APP; beta-amyloid protein; gelatinase A;
KW progelatinase A; decomposing agent; prophylaxis; Alzheimer's disease.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Region 672..711

FT /note= "beta-AP 1-40 gelatinase A
FT cleavage sites"

FT Cleavage-site 687..689

FT /label= gelatinase A

FT Cleavage-site 701..702

FT /label= gelatinase A

FT Cleavage-site 705..706

FT /label= gelatinase A

XX

PN EP622079-A2.

XX

PD 02-NOV-1994.

XX

PF 25-APR-1994; 94EP-0302924.

XX

PR 27-APR-1993; 93JP-0122207.

PR 25-FEB-1994; 94JP-0051133.

XX

PA (ORIY) ORIENTAL YEAST CO LTD.

XX

PI Miyazaki K;

XX

DR WPI; 1994-334379/42.

XX

PT Gelatinase A-contg. amyloid beta protein decomposing agent -
PT useful for prophylaxis or treatment of Alzheimer's disease

XX

PS Disclosure; Fig 1; 12pp; English.

XX

CC AAR63442 describes the amino acid sequence of the amyloid protein
CC precursor (APP), from which beta amyloid protein is derived (APB).
CC An APB decomposing agent having either gelatinase A, a limited
CC decomposate of gelatinase A, or progelatinase A as an active
CC ingredient was developed. This agent can be used in medicine for
CC prophylaxis and for the treatment of Alzheimer's disease.
CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 15; Length 770;
Best Local Similarity 100.0%; Pred. No. 9.2e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
      |||
Db     61  TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

Qy    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181  GVEFVCCPLAEESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEEEVAEVEEE 240
      |||
Db    181  GVEFVCCPLAEESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEEEVAEVEEE 240

Qy    241  EADDDDEDDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241  EADDDDEDDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC 300

Qy    301  RAMISRWYFDVTEGKCAPFFYGGCGGNRNMFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301  RAMISRWYFDVTEGKCAPFFYGGCGGNRNMFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361  PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420
      |||
Db    361  PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420

Qy    421  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480
      |||
Db    421  KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480

Qy    481  QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
      |||
Db    481  QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Qy    541  MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRI SYGNDALMPSLTET 600
      |||
Db    541  MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDVLANMISEPRI SYGNDALMPSLTET 600

Qy    601  KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
      |||
Db    601  KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Qy    661  IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAI IGLMVGGVVIATVIVITL 720
      |||
Db    661  IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAI IGLMVGGVVIATVIVITL 720

Qy    721  VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
      |||
Db    721  VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
```

RESULT 4

AAW40130

ID AAW40130 standard; Protein; 770 AA.

XX

AC AAW40130;

XX

DT 03-JUN-1998 (first entry)

XX

DE Human APP770 protein.

XX

KW Amyloid-beta peptide; A-beta; beta-AP; senile plaque; angiopathy;
 KW brain; membrane-spanning glycoprotein; beta-amyloid precursor protein;
 KW APP770; chromosome 21; human; Alzheimers disease; AD; amyloid filament;
 KW treatment; disease; Down's syndrome; hereditary cerebral haemorrhage.

XX

OS Homo sapiens.

XX

PN WO9748983-A1.

XX

PD 24-DEC-1997.

XX

PF 18-JUN-1997; 97WO-US10601.

XX

PR 18-JUN-1996; 96US-0665649.

XX

PA (ATHE-) ATHENA NEUROSCIENCES INC.

PA (BGHM) BRIGHAM & WOMENS HOSPITAL.

XX

PI Citron M, Schenk D, Selkoe DJ, Seubert PA;

XX

DR WPI; 1998-063287/06.

DR N-PSDB; AAV10322.

XX

PT Identifying compounds that alter cellular production of amyloid-beta
 PT 42 fragment - in vitro or in transgenic animal models, potentially
 PT useful for treatment of Alzheimer's and other amyloid deposition
 PT diseases

XX

PS Disclosure; Fig 10; 86pp; English.

XX

CC This sequence represents the human beta-amyloid precursor protein
 CC APP770, which is a membrane-spanning glycoprotein encoded by a gene
 CC on the long arm of chromosome 21. A fragment of the APP protein is known
 CC as the amyloid-beta peptide (A-beta), also known as the beta-AP peptide,
 CC which forms the subunit of the amyloid filaments comprising senile
 CC (amyloid) plaques and the amyloid deposits in small cerebral and
 CC meningeal blood vessels (amyloid angiopathy). The A-beta peptide can be
 CC a 39-43 amino acid fragment. This invention provides methods of screening
 CC compounds for their ability to alter the production of the A-beta
 CC peptide, which is composed of >41 amino acids, alone, or in combination
 CC with the A-beta peptide composed of 40 amino acids or less. Such agents
 CC that reduce the production of the A-beta peptide are potentially useful
 CC for treatment of Alzheimers Disease or other diseases involving amyloid
 CC deposition such as Down's syndrome, hereditary cerebral haemorrhage with
 CC amyloidosis of Dutch type and advanced aging of the brain.

XX

SQ Sequence 770 AA;

RESULT 5

AAW97996

ID AAW97996 standard; Protein; 770 AA.

XX

AC AAW97996;

XX

DT 21-JUN-1999 (first entry)

XX

DE Human amyloid precursor protein.

XX

KW Amyloid precursor protein; APP; human; gene targetting;

KW homologous recombination; transgenic mouse; transgenic animal;

KW animal model; Alzheimer's disease.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 672..711

FT /note= "beta-amyloid domain"

XX

PN WO9909150-A1.

XX

PD 25-FEB-1999.

XX

PF 18-AUG-1997; 97WO-US14507.

XX

PR 18-AUG-1997; 97WO-US14507.

XX

PA (FARB) BAYER CORP.

XX

PI Wirak DO;

XX

DR WPI; 1999-181029/15.

XX

PT Modification of target nucleic acids - by homologous recombination,
PT used particularly for introducing a humanised amyloid precursor
PT protein gene into rodents for producing models of Alzheimer's
PT disease

XX

PS Disclosure; Page 85-88; 209pp; English.

XX

CC This polypeptide comprises human amyloid precursor protein (hAPP).

CC The invention provides a novel gene targetting strategy that
CC facilitates the introduction of one or more specific mutations
CC into any gene in a single double reciprocal homologous recombination
CC step. The method has been used particularly for introducing a
CC humanised APP gene into rodents for producing animal models of
CC Alzheimer's disease (AD). 4 Independent lines of transgenic mice
CC (lines ES5007, ES5103, ES5401 and ES5403) have been created using
CC the novel gene targetting technique applied to embryonic stem cells.
CC In each line, the mouse APP gene was modified to encode a
CC mouse/human hybrid (m/hAPP) where amino acid residues 666-770 of
CC APP770 are encoded by human cDNA sequences instead of mouse genomic
CC exons (exons 16-18). Within these residues, only 3 amino acid
CC differences exist between the mouse and human proteins, i.e.
CC Gly-676 to Arg, Phe-681 to Thr and Arg-684 to His. This exon-cDNA

CC fusion gene therefore encodes an APP containing a humanised
CC beta-amyloid domain. Swedish, London, Swedish/London and stop
CC mutations have also been introduced. Targetting vector sequences
CC are provided (see AAX24730-33).

XX

SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 20; Length 770;
Best Local Similarity 100.0%; Pred. No. 9.2e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIYRCLVG 120
      |||
Db     61 TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIYRCLVG 120

Qy    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
      |||
Db    121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy    181 GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240
      |||
Db    181 GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEEE 240

Qy    241 EADDDDEDDGDEVEVEEAEEPYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300
      |||
Db    241 EADDDDEDDGDEVEVEEAEEPYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC 300

Qy    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360
      |||
Db    301 RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD 360

Qy    361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420
      |||
Db    361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA 420

Qy    421 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480
      |||
Db    421 KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480

Qy    481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
      |||
Db    481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Qy    541 MNQSLSLLYNVPAAVEEIQDEVDLLQKEQNYSDDLANMISEPRISYGNDAIMPSTLTET 600
      |||
Db    541 MNQSLSLLYNVPAAVEEIQDEVDLLQKEQNYSDDLANMISEPRISYGNDAIMPSTLTET 600

Qy    601 KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 660
      |||
Db    601 KTTVELLPVNGEFSLDDLQPWHSFGADSVANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Qy    661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
      |||
```

Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAITVIVITL 720

Qy 721 VMLKKKQYTSIHGTVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 |||||

Db 721 VMLKKKQYTSIHGTVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 6

AAE10648

ID AAE10648 standard; Protein; 770 AA.

XX

AC AAE10648;

XX

DT 10-DEC-2001 (first entry)

XX

DE Human amyloid protein precursor 770 (APP770) isoform.

XX

KW Human; aspartyl protease 1; Aspl; amyloid precursor protein 770; APP770;

KW Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;

KW amyloid plaque; neuronal loss; proteolytic; nootropic; neuroprotective.

XX

OS Homo sapiens.

XX

PN GB2357767-A.

XX

PD 04-JUL-2001.

XX

PF 22-SEP-2000; 2000GB-0023315.

XX

PR 23-SEP-1999; 99US-0155493.

PR 23-SEP-1999; 99US-0404133.

PR 23-SEP-1999; 99WO-US20881.

PR 13-OCT-1999; 99US-0416901.

PR 06-DEC-1999; 99US-0169232.

XX

PA (PHAA) PHARMACIA & UPJOHN CO.

XX

PI Bienkowski MJ, Gurney M;

XX

DR WPI; 2001-444208/48.

DR N-PSDB; AAD17897.

XX

PT Polypeptide comprising fragments of human aspartyl protease with
 PT amyloid precursor protein processing activity and alpha-secretase
 PT activity, for identifying modulators useful in treating Alzheimer's
 PT disease -

XX

PS Example 8; Page 142-144; 187pp; English.

XX

CC The patent discloses human aspartyl protease 1 (hu-Aspl) or modified
 CC Aspl proteins which lack transmembrane domain or amino terminal
 CC domain or cytoplasmic domain and retains alpha-secretase activity
 CC and amyloid protein precursor (APP) processing activity. The proteins
 CC of the invention are useful for assaying hu-Aspl alpha-secretase
 CC activity, which in turn is useful for identifying modulators of
 CC hu-Aspl alpha-secretase activity, where modulators that increase
 CC hu-Aspl alpha-secretase activity are useful for treating Alzheimer's

Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA TVIVITL 720
 |||
 Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA TVIVITL 720
 |||
 Qy 721 VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFE QMQN 770
 |||
 Db 721 VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFE QMQN 770
 |||

RESULT 7

AAE11762

ID AAE11762 standard; Protein; 770 AA.

XX

AC AAE11762;

XX

DT 18-DEC-2001 (first entry)

XX

DE Human amyloid precursor protein (APP).

XX

KW Human; neuroprotective; nootropic; immunostimulant; Alzheimer's disease;
 KW anticonvulsant; vaccine; gene therapy; Pick's disease; antidiabetic;
 KW systemic amyloidosis; maturity onset diabetes; Parkinson's disease;
 KW Huntington's disease; fronto-temporal dementia; encephalopathy; ALS;
 KW amyotrophic lateral sclerosis; amyloid precursor protein; APP.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	Peptide	1..18
FT		/label= Signal_peptide
FT	Protein	19..770
FT		/label= Mature_human_AAP_protein
FT	Domain	18..700
FT		/label= Extracellular_domain
FT	Region	672..714
FT		/note= "Abeta-42/43 core peptide"
FT	Domain	700..723
FT		/label= Transmembrane_domain
FT	Region	714..770
FT		/note= "C-100 fragment"
FT	Domain	723..770
FT		/label= Intracellular_domain

XX

PN WO200162284-A2.

XX

PD 30-AUG-2001.

XX

PF 19-FEB-2001; 2001WO-DK00113.

XX

PR 21-FEB-2000; 2000DK-0000265.

PR 01-MAR-2000; 2000US-186295P.

XX

PA (MEBI-) M & E BIOTECH AS.

XX

PI Birk P, Jensen MR, Nielsen KG;

XX

Qy	361	PVKLPPTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Db	361	PVKLPPTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGN DALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISYGN DALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEI SEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEI SEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQONGYENPTYKFFEQM QN	770
Db	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQONGYENPTYKFFEQM QN	770

111 (SERN/) CORNET M. E.

PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.
PA (YANR/) YAN R.

XX

PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;

XX

DR WPI; 2001-483072/52.

DR N-PSDB; AAD13278.

XX

PT Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity -

XX

PS Claim 8; Page 171-173; 185pp; English.

XX

CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
CC precursor protein (APP) isoforms and their corresponding DNA molecules.
CC Human aspartyl proteases can act as beta-secretase proteases useful for
CC treating Alzheimer's disease. APP isoforms are useful for identifying
CC modulators of amyloid-beta peptide production, for use in designing
CC therapeutics for the treatment and prevention of Alzheimer's disease,
CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
CC and neuronal loss. APP isoforms are also used in methods for identifying
CC inhibitors and modulators of human Asp2 activity. The invention relates
CC to a method for identifying agents that modulate the activity of human
CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
CC as a means to screen in cellular assays for the inhibitors of beta- and
CC gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in
CC polymerase chain reactions (PCR). The probes are useful for detecting
CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
CC blots. The present sequence is human wild-type amyloid precursor
CC protein 770 (APP770) isoform. APP770 gene is localised of chromosome 21.

XX

SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 22; Length 770;

Best Local Similarity 100.0%; Pred. No. 9.2e-289;

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
      |||
Db      1  MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60

Qy     61  TCIDTKEGILQYCQEVYPQLQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
      |||
Db     61  TCIDTKEGILQYCQEVYPQLQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120

Qy    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP CGIDKFR 180
      |||
Db    121  EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP CGIDKFR 180

Qy    181  GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE 240
      |||
Db    181  GVEFVCCPLAEESDNVDSADAEEEDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE 240

Qy    241  EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTSVEEVVREVCSEQAETGPC 300
```

Db	241	EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTSVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Db	361	PVKLPPTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHGHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFPEQMQN	770
Db	721	VMLKKKQYTSIHGHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFPEQMQN	770

RESULT 9

AAU06622

ID AAU06622 standard; Protein; 770 AA.

XX

AC AAU06622;

XX

DT 24-OCT-2001 (first entry)

XX

DE Human partial Amyloid precursor protein, APP770.

XX

KW Human; Aspartyl protease; Asp2; beta-secretase; nootropic;
 KW neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
 KW amyloid-beta; Abeta; APP770.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Misc-difference 358

FT /note= "Encoded by GCC"

XX

PN WO200149098-A2.

XX
PD 12-JUL-2001.
XX
PF 09-MAY-2001; 2001WO-IB00798.
XX
PR 09-MAY-2001; 2001WO-IB00798.
XX
PA (BIEN/) BIENKOWSKI M J.
PA (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.
PA (YANR/) YAN R.
XX
PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
XX
DR WPI; 2001-502549/55.
DR N-PSDB; AAS11549.
XX
PT Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity -
XX
PS Disclosure; Page 171-173; 185pp; English.
XX
CC The invention relates to a purified polypeptide comprising a fragment of
CC mammalian aspartyl protease (Asp)2 protein which lacks the Asp2
CC transmembrane domain and the Asp2 protein, and where the polypeptide and
CC the fragment retain the beta-secretase activity of the mammalian Asp2
CC protein. The invention also details polynucleotides for the Asp
CC proteins and vectors expressing them, and a polypeptide (isoform of
CC amyloid protein precursor (APP)) comprising the amino acid sequence of an
CC APP or its fragment containing an APP cleavage site recognizable by a
CC mammalian beta-secretase, and further comprising two lysine residues at
CC the carboxyl terminus of the amino acid sequence of the mammalian APP or
CC APP fragment. Also included in the invention are methods of identifying
CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
CC useful for treating Alzheimer's disease. APP is useful in methods for
CC identifying inhibitors or modulators of human Asp2 activity and
CC amyloid-beta (Abeta) peptide production. APP is also useful in designing
CC therapeutics for the treatment or prevention of Alzheimer's disease.
CC APP comprising the APP-Sw-beta-secretase peptide sequence (NLDA), which
CC is associated with increased levels of Abeta processing is useful in
CC assays relating the Alzheimer's research. The expression vector is useful
CC for recombinantly expressing APP. Nucleic acids that hybridise to
CC Asp oligonucleotides are useful as probes or primers. The probes are
CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in
CC Northern and Southern blots. The present sequence is human APP770.
XX
SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 22; Length 770;
Best Local Similarity 100.0%; Pred. No. 9.2e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK 60
|||||

Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Qy	61	TCIDTKEGILQYQCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG	120
Db	61	TCIDTKEGILQYQCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Qy	241	EADDDDEDDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Db	241	EADDDDEDDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLOKEQNYSDDVLANMISEPRI SYGNDALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDELLOKEQNYSDDVLANMISEPRI SYGNDALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMQN	770
Db	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQQMQN	770

RESULT 10

AAE02600

ID AAE02600 standard; Protein; 770 AA.

XX

AC AAE02600;

XX

DT 10-AUG-2001 (first entry)
 XX
 DE Human amyloid precursor protein 770 (APP 770) protein.
 XX
 KW Human; alpha-secretase; amyloid precursor protein; APP; therapy;
 KW Alzheimer's disease; antialzheimer's.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 358
 FT /note= "Encoded by GGC"
 XX
 PN WO200123533-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 22-SEP-2000; 2000WO-US26080.
 XX
 PR 23-SEP-1999; 99US-0155493.
 PR 23-SEP-1999; 99WO-US20881.
 PR 13-OCT-1999; 99US-0416901.
 PR 06-DEC-1999; 99US-0169232.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Gurney M, Bienkowski MJ;
 XX
 DR WPI; 2001-290516/30.
 DR N-PSDB; AAD06770.
 XX
 PT Enzymes that cleave the alpha-secretase site of the amyloid precursor
 PT protein, useful for the treatment of Alzheimer's disease -
 XX
 PS Example 8; Page 170-172; 189pp; English.
 XX
 CC The present invention relates to enzymes for cleaving the alpha-
 CC secretase site of the amyloid precursor protein (APP) and methods of
 CC identifying those enzymes. The methods may be used to identify enzymes
 CC that may be used to cleave the alpha-secretase cleavage site of the APP
 CC protein. The enzymes may be used to treat or modulate the progress of
 CC Alzheimer's disease. The present sequence is human APP 770 protein.
 XX
 SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 22; Length 770;
 Best Local Similarity 100.0%; Pred. No. 9.2e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHNMNVQNGKWDSDPSGTK 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHNMNVQNGKWDSDPSGTK 60

Qy 61 TCIDTKEGILQYQCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 TCIDTKEGILQYQCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG 120

Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Qy	241	EADDDDEDDGDEVEEEAEPEYEEATERTTSIATTTTTTTSVEEVVREVCSEQAETGPC	300
Db	241	EADDDDEDDGDEVEEEAEPEYEEATERTTSIATTTTTTTSVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Db	361	PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDPELLQKEQNYSDDLANMISEPRISYGNLALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDPELLQKEQNYSDDLANMISEPRISYGNLALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLLDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLLDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770
Db	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770

RESULT 11

ABG94279

ID ABG94279 standard; Protein; 770 AA.

XX

AC ABG94279;

XX

DT 10-DEC-2002 (first entry)

XX

DE Amyloid beta protein.

XX

KW Human; mouse; rat; antimicrobial; antiallergic; immunomodulatory;
 KW cytostatic; antiviral; antidiabetic; hypoglycaemic; antigen array;

KW vaccine; infectious disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200256905-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 21-JAN-2002; 2002WO-IB00166.
 XX
 PR 19-JAN-2001; 2001US-262379P.
 PR 04-MAY-2001; 2001US-288549P.
 PR 05-OCT-2001; 2001US-326998P.
 PR 07-NOV-2001; 2001US-331045P.
 XX
 PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
 XX
 PI Renner WA, Bachmann M, Tissot A, Maurer P, Lechner F, Sebbel P;
 PI Piossek C;
 XX
 DR WPI; 2002-627351/67.
 XX
 PT Molecular antigen array used in the production of vaccines for
 PT infectious diseases -
 XX
 PS Disclosure; Page 417-419; 441pp; English.
 XX
 CC This invention relates to a novel ordered and repetitive antigen array
 CC used in the production of vaccines for infectious diseases. The
 CC invention also discloses a composition comprising a non-natural
 CC molecular scaffold comprising a core particle selected from a core
 CC particle of a non-natural origin and a core particle of natural origin
 CC and an organiser comprising at least one first attachment site, where
 CC the organiser is connected to the core particle by at least one covalent
 CC bond. Also disclosed is an antigen or antigenic determinant with at
 CC least one second attachment site, where the antigen or antigenic
 CC determinant is amyloid beta peptide (Abeta1-42) or its fragment and
 CC where the second attachment site is selected from an attachment site not
 CC naturally occurring with the antigen or antigenic determinant and an
 CC attachment site naturally occurring with the antigen or antigenic
 CC determinant, where the second attachment site is capable of association
 CC through at least one non-peptide bond to the first attachment site and
 CC where the antigen or antigenic determinant and the scaffold interact
 CC through the association to form an ordered and repetitive antigen array.
 CC The invention also comprises a coat protein capable of forming a capsid
 CC which comprises mutant Qbeta coat proteins having an amino acid sequence
 CC selected from five amino acid sequences fully defined in the
 CC specification. The compounds of the invention may have antimicrobial,
 CC antiallergic, immunomodulatory, cytostatic, antiviral, antidiabetic,
 CC or hypoglycaemic activities and may be used in immunisation and as a
 CC vaccine. The present sequence represents a protein sequence used to
 CC create the compositions of the invention.
 XX
 SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 23; Length 770;
 Best Local Similarity 100.0%; Pred. No. 9.2e-289;

	Matches	770;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60							
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60							
Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG	120							
Db	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG	120							
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP CGIDKFR	180							
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLP CGIDKFR	180							
Qy	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240							
Db	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240							
Qy	241	EADDDDEDDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300							
Db	241	EADDDDEDDEDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300							
Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360							
Db	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360							
Qy	361	PVKLPTTAAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420							
Db	361	PVKLPTTAAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420							
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480							
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480							
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTHLRVIYER	540							
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTHLRVIYER	540							
Qy	541	MNQSLSLLYNVPAAVAEEIQDEVDELLQKEQNYSDDLANMISEPRISYGN DALMPSLTET	600							
Db	541	MNQSLSLLYNVPAAVAEEIQDEVDELLQKEQNYSDDLANMISEPRISYGN DALMPSLTET	600							
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660							
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660							
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720							
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720							
Qy	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQM QN	770							
Db	721	VMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQM QN	770							

ID ABG32723 standard; Protein; 770 AA.
XX
AC ABG32723;
XX
DT 22-NOV-2002 (first entry)
XX
DE Human amyloid precursor protein (APP) 770.
XX
KW Human; amyloid precursor protein; APP; APP695; APP770; APP751;
KW Alzheimer's disease; AD; transgenic; amyloid beta; Abeta;
KW degenerative disorder; brain; dementia; memory loss; schizophrenia;
KW neurotic plaque; cortex; hippocampus; subiculum; hippocampal gyrus;
KW amygdala; amyloid; Down's syndrome.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 289..344
FT /note= "KPI domain"
FT Domain 345..364
FT /note= "OX-2 domain"
FT Region 672..714
FT /note= "Abeta region"
XX
PN US2002104104-A1.
XX
PD 01-AUG-2002.
XX
PF 08-SEP-1998; 98US-0149718.
XX
PR 07-JUN-1995; 95US-0480653.
PR 07-JUN-1995; 95US-0486538.
PR 07-JUN-1996; 96US-0659797.
PR 07-JUN-1996; 96US-0660487.
XX
PA (GAME/) GAMES K D.
PA (SCHE/) SCHENK D B.
PA (MCCO/) MCCONLOGUE L C.
PA (SEUB/) SEUBERT P A.
PA (RYDE/) RYDEL R E.
XX
PI Games KD, Schenk DB, McConlogue LC, Seubert PA, Rydel RE;
XX
DR WPI; 2002-697836/75.
DR N-PSDB; ABS54406.
XX
PT Testing compounds for effect on Alzheimer's disease marker by using
PT transgenic mammal into which nucleic acid encoding protein including
PT APP770, APP751 or APP695 with/without specific mutations, has been
PT incorporated -
XX
PS Disclosure; Page 46-48; 62pp; English.
XX
CC The invention discloses a method for testing compounds for their effect
CC on an Alzheimer's disease (AD) marker, by administering the compound to a
CC non-human transgenic mammal which has had a nucleic acid construct stably
CC incorporated into the genome and has a promoter for expression of all,

Db 541 MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET 600

Qy 601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 |||

Db 601 KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVV IATVIVITL 720
 |||

Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVV IATVIVITL 720

Qy 721 VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFE QMQN 770
 |||

Db 721 VMLKKKQYTSIH HGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFE QMQN 770

RESULT 13

ABG80591

ID ABG80591 standard; Protein; 770 AA.

XX

AC ABG80591;

XX

DT 29-NOV-2002 (first entry)

XX

DE Human amyloid beta protein.

XX

KW Molecular antigen array; vaccine; antigen; antimicrobial;

KW molecular scaffold; amyloid beta; Abeta 1-42; influenza;

KW graft versus host disease; IgE-mediated allergic reaction; anaphylaxis;

KW adult respiratory distress syndrome; ARDS; Crohn's disease;

KW allergic asthma; acute lymphoblastic leukaemia; non-Hodgkin's lymphoma;

KW Grave's disease; systemic lupus erythematosus; osteoporosis;

KW inflammatory immune disease; myasthenia gravis; multiple sclerosis;

KW immunoproliferative disease lymphadenopathy; Alzheimer's disease;

KW angioimmunoproliferative lymphadenopathy; immunoblastic lymphadenopathy;

KW rheumatoid arthritis; diabetes; infectious disease; factor Xa;

KW enterokinase; cysteine-containing linker.

XX

OS Homo sapiens.

XX

PN WO200256907-A2.

XX

PD 25-JUL-2002.

XX

PF 21-JAN-2002; 2002WO-IB00168.

XX

PR 19-JAN-2001; 2001US-262379P.

PR 04-MAY-2001; 2001US-288549P.

PR 05-OCT-2001; 2001US-326998P.

PR 07-NOV-2001; 2001US-331045P.

XX

PA (CYTO-) CYTOS BIOTECHNOLOGY AG.

PA (NOVS) NOVARTIS PHARMA AG.

PA (MAUR/) MAURER P.

PA (LECH/) LECHNER F.

PA (ORTM/) ORTMANN R.

PA (LUEO/) LUEOEND R.

PA (STAU/) STAUFENBIEL M.

Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEEDSDVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Qy	241	EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Db	241	EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTTESVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPVAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDAIMPSTET	600
Db	541	MNQSLSLLYNVPVAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDAIMPSTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFVFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFVFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770
Db	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770

RESULT 14

ABG76936

ID ABG76936 standard; Protein; 770 AA.

XX

AC ABG76936;

XX

DT 05-NOV-2002 (first entry)

XX

DE Humanised antibody associated protein #5.

XX

KW Humanized; immunoglobulin; Ig; light chain; LC; heavy chain; HC;

KW variable region complementarity determining region; 3D6; 10D5;
KW variable framework region; amyloidogenic disease; Alzheimer's disease;
KW amyloid deposit; variable light chain; VL; variable heavy chain; VH;
KW nootropic; neuroprotective; inhibitor of beta amyloid accumulation;
KW Abeta; antibody.
XX
OS Homo sapiens.
XX
PN WO200246237-A2.
XX
PD 13-JUN-2002.
XX
PF 06-DEC-2001; 2001WO-US46587.
XX
PR 06-DEC-2000; 2000US-251892P.
XX
PA (NEUR-) NEURALAB LTD.
PA (AMHP) WYETH.
XX
PI Basi G, Saldanha J, Yednock T;
XX
DR WPI; 2002-519658/55.
XX
PT Novel light/heavy chain of humanized immunoglobulin for treating
PT amyloidogenic disease, has 3D6/10D5 variable region complementarity
PT determining regions and variable framework region from human acceptor
PT immunoglobulin -
XX
PS Disclosure; Page 165-167; 171pp; English.
XX
CC The present invention relates to new humanized immunoglobulin (Ig) light
CC chain (LC) or heavy chain (HC) comprising variable region complementarity
CC determining regions from 3D6/10D5 Ig LC or HC variable region sequence,
CC and variable framework region from human acceptor Ig LC or HC sequence.
CC The invention is useful for preventing or treating an amyloidogenic
CC disease or Alzheimer's disease in a patient. The invention is also useful
CC for in vivo imaging amyloid deposits in a patient. The present amino
CC acid sequence represents a humanized antibody associated protein as
CC described in the invention.
XX
SQ Sequence. 770 AA;

Query Match 100.0%; Score 4058; DB 23; Length 770;
Best Local Similarity 100.0%; Pred. No. 9.2e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIRCLVG	120
Db	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPIRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180

Qy	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Qy	241	EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC	300
Db	241	EADDDDEDDGDEVEEEAEPEYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTTEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTTEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPVAVAEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET	600
Db	541	MNQSLSLLYNVPVAVAEEIQDEVDELLQKEQNYSDVLANMISEPRISYGNDALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLLDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLLDLQPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770
Db	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	770

RESULT 15

ABB78609

ID ABB78609 standard; Protein; 770 AA.

XX

AC ABB78609;

XX

DT 16-JUL-2002 (first entry)

XX

DE Human APP770 protein sequence SEQ ID NO:55.

XX

KW Human; Asp-1; Asp-2; aspartyl protease; Alzheimer's disease;
 KW proteolytic; amyloid precursor protein; APP.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers
 FT Misc-difference 358
 FT /note= "encoded by GGC"
 XX
 PN GB2367060-A.
 XX
 PD 27-MAR-2002.
 XX
 PF 29-OCT-2001; 2001GB-0025934.
 XX
 PR 23-SEP-1999; 99US-155493P.
 PR 23-SEP-1999; 99US-0404133.
 PR 23-SEP-1999; 99WO-US20881.
 PR 13-OCT-1999; 99US-0416901.
 PR 06-DEC-1999; 99US-169232P.
 PR 22-SEP-2000; 2000GB-0023315.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Bienkowski MJ, Gurney M;
 XX
 DR WPI; 2002-396337/43.
 DR N-PSDB; ABL52489.
 XX
 PT Human aspartyl protease 1 substrates useful in assays to detect
 PT aspartyl protease activity, e.g. for the diagnosis of Alzheimer's
 PT disease -
 XX
 PS Disclosure; Page 142-144; 182pp; English.
 XX
 CC The present invention describes a human aspartyl protease 1 (hu-Aspl)
 CC substrate (I) which comprises a peptide of no more than 50 amino acids,
 CC and which comprises the 8 amino acid sequence Gly-Leu-Ala-Leu-Ala-Leu-
 CC Glu-Pro. Also described are: (1) a method (II) for assaying hu-Aspl
 CC proteolytic activity, comprising: (a) contacting a hu-Aspl protein with
 CC (I) under acidic conditions; and (b) determining the level of hu-Aspl
 CC proteolytic activity; (2) a purified polynucleotide (III) comprising a
 CC nucleotide sequence that hybridises under stringent conditions to the
 CC non-coding strand complementary to a defined 1804 nucleotide sequence
 CC (see ABL52456) where the nucleotide sequence encodes a polypeptide having
 CC Aspl proteolytic activity and lacks nucleotides encoding a transmembrane
 CC domain); (3) a purified polynucleotide (III') comprising a sequence that
 CC hybridises under stringent conditions to (III) (the nucleotide sequence
 CC encodes a polypeptide further lacking a pro-peptide domain corresponding
 CC to amino acids 23-62 of hu-Aspl (see ABB78589)); (4) a vector (IV)
 CC comprising (III) or (III'); and (5) a host cell (V) transformed or
 CC transfected with (III), (III') and/or (IV). The hu-Aspl protease
 CC substrate (I) may be used as an enzyme substrate in assays to detect
 CC aspartyl protease activity, (II) and therefore diagnose diseases
 CC associated with aberrant hu-Aspl expression and activity such as
 CC Alzheimer's disease. Hu-Aspl has been localised to chromosome 21, while
 CC hu-Asp2 has been localised to chromosome 11q23.3-24.1. The present
 CC sequence represents human amyloid precursor protein APP770, which is used
 CC in the exemplification of the present invention.
 XX
 SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 23; Length 770;
Best Local Similarity 100.0%; Pred. No. 9.2e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK	60
Qy	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG	120
Db	61	TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVI PYRCLVG	120
Qy	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR	180
Qy	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Db	181	GVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVAEEEEVAEVEEE	240
Qy	241	EADDDDEDDGDEVEEEAEFPYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC	300
Db	241	EADDDDEDDGDEVEEEAEFPYEEATERTTTSIATTTTTTTESVEEVVREVCSEQAETGPC	300
Qy	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Db	301	RAMISRWFYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMVCGSAMSQSLLKTTQEPLARD	360
Qy	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Db	361	PVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA	420
Qy	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Db	421	KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL	480
Qy	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Db	481	QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER	540
Qy	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET	600
Db	541	MNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDDLANMISEPRISYGNDALMPSLTET	600
Qy	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Db	601	KTTVELLPVNGEFSLDDLQPWHSFGADSV PANTENEVEPVDARPAADRGLTTRPGSGLTN	660
Qy	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL	720
Qy	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQM QN	770
Db	721	VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQM QN	770

Search completed: December 13, 2003, 13:34:22
Job time : 130.863 secs

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 2.0696 Seconds
(without alignments)
1870.310 Million cell updates/sec

Title: US-09-785-215-4
Perfect score: 74
Sequence: 1 QYIKANSKFIGITEL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							

1	74	100.0	1310	2	Q93N27	Q93n27 clostridium
2	44.5	60.1	60	7	Q31585	Q31585 salmo salar
3	44.5	60.1	71	7	Q9XRJ9	Q9xrxj9 salvelinus
4	44.5	60.1	85	7	Q95IS2	Q95is2 salmo salar
5	44.5	60.1	85	7	Q95HY1	Q95hyl1 salmo salar
6	44.5	60.1	86	7	Q95HX4	Q95hx4 salmo salar
7	44.5	60.1	244	7	Q31590	Q31590 salmo salar
8	44	59.5	546	10	Q9XG37	Q9xg37 guillardia
9	43.5	58.8	67	7	Q31578	Q31578 salmo salar
10	43	58.1	180	16	Q9CF66	Q9cf66 lactococcus
11	43	58.1	250	9	Q9MCL7	Q9mcl7 streptococc
12	43	58.1	252	9	Q9XJE8	Q9xje8 lactococcus
13	43	58.1	291	11	Q9CRV4	Q9crv4 mus musculu
14	43	58.1	304	11	Q8K2A1	Q8k2a1 mus musculu
15	43	58.1	309	11	Q9CYD2	Q9cyd2 mus musculu
16	43	58.1	572	10	Q8H8F3	Q8h8f3 oryza sativ
17	43	58.1	899	12	Q9YTK4	Q9ytk4 ateline her
18	42.5	57.4	1087	5	Q25961	Q25961 plasmodium
19	42.5	57.4	1694	5	Q9NHX1	Q9nhx1 plasmodium
20	42.5	57.4	1694	5	Q9Tzt5	Q9tzt5 plasmodium
21	42.5	57.4	1704	5	Q9Tzt4	Q9tzt4 plasmodium
22	42.5	57.4	1720	5	Q25922	Q25922 plasmodium
23	42.5	57.4	1720	5	Q8I0U8	Q8i0u8 plasmodium
24	42	56.8	358	17	Q8Tzu8	Q8tzu8 pyrococcus
25	42	56.8	1333	5	Q24262	Q24262 drosophila
26	41.5	56.1	84	13	Q9DEK4	Q9dek4 coregonus s
27	41.5	56.1	85	7	Q95IS3	Q95is3 salmo salar
28	41.5	56.1	85	7	Q95IR2	Q95ir2 salmo salar
29	41.5	56.1	149	7	Q31495	Q31495 oncorhynch
30	41.5	56.1	216	7	Q9GJH0	Q9gjh0 salmo trutt
31	41.5	56.1	216	7	Q9GJG9	Q9gjjg9 salmo trutt
32	41	55.4	187	16	Q8R9Q7	Q8r9q7 thermoanaer
33	41	55.4	247	11	Q9D3B9	Q9d3b9 mus musculu
34	41	55.4	384	4	Q9HD07	Q9hd07 homo sapien
35	41	55.4	447	16	Q97HK7	Q97hk7 clostridium
36	41	55.4	532	5	Q96671	Q96671 drosophila
37	41	55.4	532	5	Q9VU53	Q9vu53 drosophila
38	41	55.4	595	11	Q8CAD2	Q8cad2 mus musculu
39	41	55.4	626	11	Q8BRL1	Q8brl1 mus musculu
40	41	55.4	689	5	Q8IHZ2	Q8ihz2 plasmodium
41	41	55.4	749	4	Q8NEI0	Q8nei0 homo sapien
42	40.5	54.7	67	7	Q31582	Q31582 salmo salar
43	40.5	54.7	67	7	Q31577	Q31577 salmo salar
44	40.5	54.7	67	7	Q31581	Q31581 salmo salar
45	40.5	54.7	71	7	Q9XRH6	Q9xrh6 salvelinus

ALIGNMENTS

RESULT 1

Q93N27

ID Q93N27 PRELIMINARY; PRT; 1310 AA.

AC Q93N27;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Tetanus toxin (Fragment).
 OS Clostridium tetani.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shumin Z., Dianliang L.;
 RT "Cloning and sequence analysis of tetanus toxin gene."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF389424; AAK72964.2; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR001064; Crystallin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 FT NON_TER 1 1
 FT NON_TER 1310 1310
 SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;

Query Match 100.0%; Score 74; DB 2; Length 1310;
 Best Local Similarity 100.0%; Pred. No. 0.00022;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||
 Db 831 QYIKANSKFIGITEL 845

RESULT 2

Q31585

ID Q31585 PRELIMINARY; PRT; 60 AA.
 AC Q31585;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE (DB02) MHC class II beta 1 (Fragment).
 OS Salmo salar (Atlantic salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 OX NCBI_TaxID=8030;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Grimholt U., Olsaker I., de Vries Lindstrom C., Lie O.;
 RT "A study of polymorphism in the MHC class II beta 1 and MHC class I
 alpha 2 domain exons of Atlantic salmon (Salmo salar).";
 RL Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; L24953; AAA49597.1; -.
 DR InterPro; IPR000353; MHC_II_beta.
 DR Pfam; PF00969; MHC_II_beta; 1.
 DR ProDom; PD000328; MHC_II_beta; 1.
 KW Glycoprotein; MHC II; Transmembrane.
 FT NON_TER 1 1

FT NON_TER 60 60
SQ SEQUENCE 60 AA; 6776 MW; 5D4F3449060940E2 CRC64;

Query Match 60.1%; Score 44.5; DB 7; Length 60;
Best Local Similarity 55.6%; Pred. No. 2;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 1 QYIKANS---KFIGITEL 15
: ||: || ||: || ||
Db 16 EYIRFNSTVGKFGVGYTEL 33

RESULT 3

Q9XRJ9

ID Q9XRJ9 PRELIMINARY; PRT; 71 AA.
AC Q9XRJ9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC class II beta 1 (Fragment).
GN SANA.
OS Salvelinus namaycush (lake trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8040;
RN [1]
RP SEQUENCE FROM N.A.
RA Dorschner M.O., Duris T., Phillips R.B.;
RT "Diversity of a Lake Trout Mhc class II Gene.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF130026; AAD20889.1; -.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1 1
FT NON_TER 71 71
SQ SEQUENCE 71 AA; 8002 MW; E3095286582A9F2D CRC64;

Query Match 60.1%; Score 44.5; DB 7; Length 71;
Best Local Similarity 55.6%; Pred. No. 2.3;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 1 QYIKANS---KFIGITEL 15
: ||: || ||: || ||
Db 14 EYIRFNSTVGKFGVGYTEL 31

RESULT 4

Q95IS2

ID Q95IS2 PRELIMINARY; PRT; 85 AA.
AC Q95IS2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE MHC class II beta chain (Fragment).

OS Salmo salar (Atlantic salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 OX NCBI_TaxID=8030;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Landry C., Bernatchez L.;
 RT "Comparative analysis of population structure across environments and
 RT geographic scales at Major Histocompatibility Complex and
 RT microsatellite in Atlantic salmon (Salmo salar).";
 RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF373699; AAK61882.1; -.
 DR InterPro; IPR000353; MHC_II_beta.
 DR Pfam; PF00969; MHC_II_beta; 1.
 DR ProDom; PD000328; MHC_II_beta; 1.
 KW Glycoprotein; MHC; MHC II; Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 85 85
 SQ SEQUENCE 85 AA; 9743 MW; 3214E01AD1B66AC5 CRC64;

 Query Match 60.1%; Score 44.5; DB 7; Length 85;
 Best Local Similarity 55.6%; Pred. No. 2.8;
 Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

 Qy 1 QYIKANS---KFIGITEL 15
 :||: || ||:| |||
 Db 33 EYIRFNSTVGKFGVGYTEL 50

RESULT 5

Q95HY1
 ID Q95HY1 PRELIMINARY; PRT; 85 AA.
 AC Q95HY1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE MHC class II B antigen (Fragment).
 GN DB1.
 OS Salmo salar (Atlantic salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 OX NCBI_TaxID=8030;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21383619; PubMed=11491536;
 RA Langefors A., Lohm J., von Schantz T.;
 RT "Allelic polymorphism in MHC class II B in four populations of
 RT Atlantic salmon (Salmo salar).";
 RL Immunogenetics 53:329-336(2001).
 DR EMBL; AF104370; AAL04002.1; -.
 DR InterPro; IPR000353; MHC_II_beta.
 DR Pfam; PF00969; MHC_II_beta; 1.
 DR ProDom; PD000328; MHC_II_beta; 1.
 KW Glycoprotein; MHC; MHC II; Transmembrane.
 FT NON_TER 1 1

FT NON_TER 85 85
SQ SEQUENCE 85 AA; 9723 MW; 27C9F7931F1F01C4 CRC64;

Query Match 60.1%; Score 44.5; DB 7; Length 85;
Best Local Similarity 55.6%; Pred. No. 2.8;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 1 QYIKANS---KFIGITEL 15
:|:| | | | | | | | |
Db 33 EYIRFNSTVGKFGVGYTEL 50

RESULT 6

Q95HX4

ID Q95HX4 PRELIMINARY; PRT; 86 AA.
AC Q95HX4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE MHC class II B antigen (Fragment).
GN DB1.
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21383619; PubMed=11491536;
RA Langefors A., Lohm J., von Schantz T.;
RT "Allelic polymorphism in MHC class II B in four populations of
RT Atlantic salmon (Salmo salar).";
RL Immunogenetics 53:329-336(2001).
DR EMBL; AF104377; AAL04009.1; -.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC; MHC II; Transmembrane.
FT NON_TER 1 1
FT NON_TER 86 86
SQ SEQUENCE 86 AA; 9912 MW; E5097729F681F149 CRC64;

Query Match 60.1%; Score 44.5; DB 7; Length 86;
Best Local Similarity 55.6%; Pred. No. 2.8;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 1 QYIKANS---KFIGITEL 15
:|:| | | | | | | | |
Db 33 EYIRFNSTVGKFGVGYTEL 50

RESULT 7

Q31590

ID Q31590 PRELIMINARY; PRT; 244 AA.
AC Q31590;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE MHC class II.
 GN MHC-SASA CLASS II B.
 OS Salmo salar (Atlantic salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 OX NCBI_TaxID=8030;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leukocyte;
 RX MEDLINE=93170890; PubMed=8436418;
 RA Hordvik I., Grimholt U., Fosse V.M., Lie Y., Endresen C.;
 RT "Cloning and sequence analysis of cDNAs encoding the MHC class II a-
 RT chain in Atlantic salmon, Salmo salar.";
 RL Immunogenetics 37:437-441(1993).
 DR EMBL; X70166; CAA49725.1; -.
 DR HSSP; P01888; 1BMG.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003597; Ig_c1.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR000353; MHC_II_beta.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00969; MHC_II_beta; 1.
 DR ProDom; PD000328; MHC_II_beta; 1.
 DR SMART; SM00407; IGc1; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Glycoprotein; MHC II; Transmembrane.
 SQ SEQUENCE 244 AA; 27449 MW; 496CB9EA9D73765C CRC64;

Query Match 60.1%; Score 44.5; DB 7; Length 244;
 Best Local Similarity 55.6%; Pred. No. 7.9;
 Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 1 QYIKANS---KFIGITEL 15
 :||: || ||:| |||
 Db 51 EYIRFNSTVGKFGVGYTEL 68

RESULT 8

Q9XG37

ID Q9XG37 PRELIMINARY; PRT; 546 AA.
 AC Q9XG37;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 66.2 kDa protein.
 OS Guillardia theta (Cryptomonas phi).
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
 OX NCBI_TaxID=55529;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20087226; PubMed=10618395;
 RA Zauner S., Fraunholz M., Wastl J., Penny S.L., Beaton M.,
 RA Cavalier-Smith T., Maier U., Douglas S.;
 RT "Chloroplast protein and centrosomal genes, a tRNA intron, and odd
 RT telomeres in an unusually compact eukaryotic genome, the cryptomonad

RT nucleomorph.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21223349; PubMed=11323671;
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.,
 RA Wu X., Reith M., Cavalier-Smith T., Maier U.;
 RT "The highly reduced genome of an enslaved algal nucleus.";
 RL Nature 410:1091-1096(2001).
 DR EMBL; AJ010592; CAB40403.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 546 AA; 66218 MW; 7303950F632BE6F2 CRC64;

Query Match 59.5%; Score 44; DB 10; Length 546;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YIKANSKFIGITEL 15
 :||:|:|:|:|:|:
 Db 445 FIKSNSRFMRLTEI 458

RESULT 9

Q31578

ID Q31578 PRELIMINARY; PRT; 67 AA.
 AC Q31578;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE (DB03) MHC class II beta 1 (Fragment).
 OS Salmo salar (Atlantic salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 OX NCBI_TaxID=8030;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Grimholt U., Olsaker I., de Vries Lindstrom C., Lie O.;
 RT "A study of polymorphism in the MHC class II beta 1 and MHC class I
 RT alpha 2 domain exons of Atlantic salmon (Salmo salar).";
 RL Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; L24929; AAA49590.1; -.
 DR InterPro; IPR000353; MHC_II_beta.
 DR Pfam; PF00969; MHC_II_beta; 1.
 DR ProDom; PD000328; MHC_II_beta; 1.
 KW Glycoprotein; MHC II; Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 67 67
 SQ SEQUENCE 67 AA; 7449 MW; 42771AEDBABA6626 CRC64;

Query Match 58.8%; Score 43.5; DB 7; Length 67;
 Best Local Similarity 50.0%; Pred. No. 3.3;
 Matches 9; Conservative 4; Mismatches 2; Indels 3; Gaps 1;

Qy 1 QYIKANS---KFIGITEL 15
 :|::||| ||:| |||
 Db 16 EYVRFNSTVVGKFGVGYTEL 33

RESULT 10

Q9CF66

ID Q9CF66 PRELIMINARY; PRT; 180 AA.
AC Q9CF66;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Spermidine acetyltransferase (EC 2.3.1.57).
GN YQFF OR LL1615.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006391; AAK05713.1; -.
DR InterPro; IPR000182; GCN5acetyltransf.
DR Pfam; PF00583; Acetyltransf; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 180 AA; 21022 MW; 6DBD148524C0DF3C CRC64;

Query Match 58.1%; Score 43; DB 16; Length 180;
Best Local Similarity 69.2%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 IKANSKFIGITEL 15
|:| | ||| |
Db 65 IEANDTFIGIVEL 77

RESULT 11

Q9MCL7

ID Q9MCL7 PRELIMINARY; PRT; 250 AA.
AC Q9MCL7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ORF13.
GN ORF13.
OS Streptococcus thermophilus bacteriophage 7201.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=112023;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20088830; PubMed=10620678;
RA Stanley E., Walsh L., van der Zwet A., Fitzgerald G.F.,
RA van Sinderen D.;
RT "Identification of four loci isolated from two Streptococcus
RT thermophilus phage genomes responsible for mediating bacteriophage

RT resistance.";
 RL FEMS Microbiol. Lett. 182:271-277(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Stanley E., Walsh L., Fitzgerald G.F., van Sinderen D.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF145054; AAF43506.1; -.
 DR InterPro; IPR001091; CN4_Mettransf.
 DR InterPro; IPR002295; D21N6_mtfrase.
 DR InterPro; IPR002941; N6/N4_Mtase.
 DR Pfam; PF01555; N6_N4_Mtase; 1.
 DR PRINTS; PR00506; D21N6MTFRASE.
 DR PRINTS; PR00508; S21N4MTFRASE.
 SQ SEQUENCE 250 AA; 28744 MW; 725EA3B2B56B0D7F CRC64;

Query Match 58.1%; Score 43; DB 9; Length 250;
 Best Local Similarity 57.1%; Pred. No. 15;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14
 | :||| | :| ||
 Db 131 QVLKANMKIVGATE 144

RESULT 12

Q9XJE8

ID Q9XJE8 PRELIMINARY; PRT; 252 AA.
 AC Q9XJE8;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative methylase.
 OS Lactococcus lactis bacteriophage Tuc2009.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
 OX NCBI_TaxID=35241;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA van Sinderen D., van de Guchte M., Seegers J.F.M.L., Fitzgerald G.F.;
 RT "Molecular analysis of the temperate lactococcal phage Tuc2009.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF109874; AAD37103.1; -.
 DR InterPro; IPR001091; CN4_Mettransf.
 DR InterPro; IPR002295; D21N6_mtfrase.
 DR InterPro; IPR002941; N6/N4_Mtase.
 DR Pfam; PF01555; N6_N4_Mtase; 1.
 DR PRINTS; PR00506; D21N6MTFRASE.
 DR PRINTS; PR00508; S21N4MTFRASE.
 KW Methyltransferase.
 SQ SEQUENCE 252 AA; 29357 MW; 4D5A59FF47363948 CRC64;

Query Match 58.1%; Score 43; DB 9; Length 252;
 Best Local Similarity 57.1%; Pred. No. 15;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14
 | :||| | :| ||
 Db 130 QVLKANMKIVGATE 143

RESULT 13

Q9CRV4

ID Q9CRV4 PRELIMINARY; PRT; 291 AA.
AC Q9CRV4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 13 days embryo head cDNA, RIKEN full-length enriched library,
DE clone:3110030A04 product:CED-6 PROTEIN homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20499374; PubMed=11042159;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 DR EMBL; AK014093; BAB29151.2; -.
 SQ SEQUENCE 291 AA; 32802 MW; 91FF3F41CD0BC6F2 CRC64;

Query Match 58.1%; Score 43; DB 11; Length 291;
 Best Local Similarity 57.1%; Pred. No. 17;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 YIKANSKFIGITEL 15
 || |:|:| ||:
 Db 11 YIPYNAKFLGSTEV 24

RESULT 14

Q8K2A1

ID Q8K2A1 PRELIMINARY; PRT; 304 AA.
 AC Q8K2A1;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to CED-6 protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; BC032154; AAH32154.1; -.
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR InterPro; IPR006020; PTB_PID.
 DR Pfam; PF00640; PID; 1.
 DR SMART; SM00462; PTB; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
 DR PROSITE; PS01179; PID; 1.
 SQ SEQUENCE 304 AA; 34470 MW; D99154EF53EFDC45 CRC64;

Query Match 58.1%; Score 43; DB 11; Length 304;

Best Local Similarity 57.1%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 YIKANSKFIGITEL 15
|| |:||:| ||:
Db 24 YIPYNAKFLGSTEV 37

RESULT 15

Q9CYD2

ID Q9CYD2 PRELIMINARY; PRT; 309 AA.
AC Q9CYD2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 5730529006Rik protein.
GN CED6 OR 5730529006RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK017798; BAB30939.1; -.
DR HSSP; Q02410; 1AQC.
DR MGD; MGI:1920407; Ced6.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR006020; PTB_PID.
DR Pfam; PF00640; PID; 1.
DR SMART; SM00462; PTB; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
DR PROSITE; PS01179; PID; 1.
SQ SEQUENCE 309 AA; 35272 MW; 535DD8E733C0F406 CRC64;

Query Match 58.1%; Score 43; DB 11; Length 309;

Best Local Similarity 57.1%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 YIKANSKFIGITEL 15
|| |:||:| ||:
Db 24 YIPYNAKFLGSTEV 37

Search completed: December 13, 2003, 13:29:50
Job time : 4.0696 secs

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 0.512821 Seconds
(without alignments)
1375.531 Million cell updates/sec

Title: US-09-785-215-4
Perfect score: 74
Sequence: 1 QYIKANSKFIGITEL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	74	100.0	1314	1	TETX_CLOTE	P04958 clostridium
2	44	59.5	66	1	VG84_BPML5	Q05301 mycobacteri
3	42.5	57.4	1682	1	MSP1_PLAF3	P19598 plasmodium
4	42.5	57.4	1701	1	MSP1_PLAFF	P13819 plasmodium
5	42.5	57.4	1701	1	MSP1_PLAFM	P08569 plasmodium
6	42.5	57.4	1726	1	MSP1_PLAFC	P04934 plasmodium
7	42.5	57.4	1726	1	MSP1_PLAFP	P50495 plasmodium
8	41	55.4	204	1	PYRC_SERMA	Q9s3s1 serratia ma
9	41	55.4	384	1	LE12_THETN	Q8rcf9 thermoanaer
10	41	55.4	899	1	V120_HSVSA	Q01055 herpesvirus
11	40	54.1	194	1	ACPD_HAEIN	P43013 haemophilus
12	40	54.1	601	1	PEF1_LACLC	P54124 lactococcus
13	40	54.1	601	1	PEPF_LACLA	Q9cev7 lactococcus
14	40	54.1	644	1	YHJ9_YEAST	P38694 saccharomyc
15	39	52.7	213	1	KAD_MYCCA	P10251 mycoplasma
16	39	52.7	505	1	GPMI_MYCPU	Q98qa7 mycoplasma
17	38	51.4	256	1	YD83_METJA	Q58778 methanococc

18	38	51.4	287	1	TRUB_AQUAE	O66922	aquifex aeo
19	38	51.4	572	1	HEMA_PI3HT	P12562	human parai
20	37	50.0	191	1	Y096_HAEIN	P43940	haemophilus
21	37	50.0	231	1	PYRF_MYCPE	Q8euy3	mycoplasma
22	37	50.0	445	1	GNT1_HUMAN	P26572	homo sapien
23	37	50.0	447	1	GNT1_MOUSE	P27808	mus musculu
24	37	50.0	447	1	GNT1_RABIT	P27115	oryctolagus
25	37	50.0	447	1	GNT1_RAT	Q09325	rattus norv
26	37	50.0	490	1	Y032_BORBU	O51063	borrelia bu
27	37	50.0	510	1	G6PD_ASPNG	P48826	aspergillus
28	37	50.0	511	1	G6PD_EMENI	P41764	emericella
29	37	50.0	548	1	2AD1_SCHPO	Q10428	schizosacch
30	36	48.6	169	1	Y358_BUCAI	P57439	buchnera ap
31	36	48.6	258	1	MIP_CHLPN	Q9z7p3	chlamydia p
32	36	48.6	296	1	YD01_CLOAB	P33659	clostridium
33	36	48.6	333	1	DPOB_XENLA	O57383	xenopus lae
34	36	48.6	451	1	MURD_BACSU	Q03522	bacillus su
35	36	48.6	461	1	NIFN_RHOCA	P19077	rhodobacter
36	36	48.6	495	1	G6PD_PICJA	P11410	pichia jadi
37	36	48.6	627	1	2AD2_SCHPO	P78759	schizosacch
38	36	48.6	630	1	YND1_YEAST	P40009	saccharomyc
39	36	48.6	774	1	RRP3_INCBE	P21770	influenza c
40	36	48.6	774	1	RRP3_INCJJ	P13877	influenza c
41	36	48.6	1630	1	MSP1_PLAFK	P04932	plasmodium
42	36	48.6	1639	1	MSP1_PLAFW	P04933	plasmodium
43	35	47.3	176	1	NU6C_SPIOL	Q9m3i8	spinacia ol
44	35	47.3	200	1	ACD2_CLOAB	Q97dql	clostridium
45	35	47.3	200	1	ACPD_SALTY	Q8xftp4	salmonella

ALIGNMENTS

RESULT 1

TETX_CLOTE

ID TETX_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) [Contains:
DE Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy
DE chain (Tetanus toxin chain H)].
GN TETX OR CTP60.
OS Clostridium tetani.
OG Plasmid pE88, and Plasmid 75 Kbp.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudel M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins."

RL EMBO J. 5:2495-2502(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CN3911; PLASMID=75 Kbp;
 RX MEDLINE=87040747; PubMed=3774547;
 RA Fairweather N.F., Lyness V.A.;
 RT "The complete nucleotide sequence of tetanus toxin.";
 RL Nucleic Acids Res. 14:7809-7812(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Massachusetts / E88; PLASMID=pE88;
 RX MEDLINE=22457253; PubMed=12552129;
 RA Brueggemann H., Baeumer S., Fricke W.F., Wiezer A., Liesegang H.,
 RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
 RA Gottschalk G.;
 RT "The genome sequence of Clostridium tetani, the causative agent of
 RT tetanus disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
 RN [4]
 RP SEQUENCE OF 742-1314 FROM N.A.
 RC PLASMID=75 Kbp;
 RX MEDLINE=86085672; PubMed=3510187;
 RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
 RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
 RT fragment C in Escherichia coli.";
 RL J. Bacteriol. 165:21-27(1986).
 RN [5]
 RP PARTIAL SEQUENCE; AND DISULFIDE BONDS.
 RX MEDLINE=90201034; PubMed=2108021;
 RA Krieglstein K., Henschen A., Weller U., Habermann E.;
 RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
 RT in tetanus toxin.";
 RL Eur. J. Biochem. 188:39-45(1990).
 RN [6]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=92037649; PubMed=1935979;
 RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
 RT "Limited proteolysis of tetanus toxin. Relation to activity and
 RT identification of cleavage sites.";
 RL Eur. J. Biochem. 202:41-51(1991).
 RN [7]
 RP IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93010948; PubMed=1396558;
 RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
 RA Montecucco C.;
 RT "Tetanus toxin is a zinc protein and its inhibition of
 RT neurotransmitter release and protease activity depend on zinc.";
 RL EMBO J. 11:3577-3583(1992).
 RN [8]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
 RT by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 RN [9]

RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
 RX MEDLINE=97475217; PubMed=9334741;
 RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
 RA Sax M.;
 RT "Structure of the receptor binding fragment HC of tetanus
 RT neurotoxin.";
 RL Nat. Struct. Biol. 4:788-792(1997).
 CC -!- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
 CC BOND OF SYNAPTOBREVIN-2.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in
 CC synaptobrevin 2.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
 CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
 CC AND ARE NON-TOXIC AFTER SEPARATION.
 CC -!- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
 CC GANGLIOSIDE RECEPTORS.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X04436; CAA28033.1; -.
 DR EMBL; X06214; CAA29564.1; -.
 DR EMBL; AF528097; AA037454.1; -.
 DR EMBL; M12739; AAA23282.1; -.
 DR PIR; A25689; BTCLTN.
 DR PDB; 1AF9; 29-APR-98.
 DR PDB; 1A8D; 14-OCT-98.
 DR PDB; 1D0H; 27-MAR-00.
 DR PDB; 1DFQ; 24-MAR-00.
 DR PDB; 1DIW; 24-MAR-00.
 DR PDB; 1DLL; 24-MAR-00.
 DR PDB; 1FV3; 05-SEP-01.
 DR MEROPS; M27.001; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_Mtpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
 KW 3D-structure; Complete proteome.
 FT INIT_MET 0 0
 FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.
 FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.
 FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).

FT	ACT_SITE	233	233
FT	METAL	236	236
FT	TRANSMEM	226	246
FT	TRANSMEM	669	689
FT	DISULFID	438	466
FT	DISULFID	1076	1092
FT	HELIX	876	882
FT	TURN	883	883
FT	STRAND	884	891
FT	TURN	892	893
FT	STRAND	894	897
FT	STRAND	904	907
FT	TURN	909	910
FT	STRAND	912	915
FT	STRAND	920	925
FT	TURN	928	929
FT	STRAND	932	935
FT	HELIX	938	940
FT	TURN	941	946
FT	STRAND	949	956
FT	HELIX	962	968
FT	TURN	969	970
FT	STRAND	972	977
FT	STRAND	980	981
FT	HELIX	983	985
FT	STRAND	987	995
FT	TURN	996	997
FT	STRAND	998	1004
FT	TURN	1006	1007
FT	STRAND	1010	1016
FT	STRAND	1020	1020
FT	TURN	1021	1022
FT	STRAND	1031	1037
FT	TURN	1039	1040
FT	STRAND	1042	1047
FT	TURN	1048	1049
FT	STRAND	1050	1056
FT	TURN	1058	1059
FT	STRAND	1068	1074
FT	TURN	1079	1080
FT	STRAND	1082	1091
FT	HELIX	1097	1105
FT	TURN	1106	1107
FT	STRAND	1112	1112
FT	STRAND	1114	1114
FT	TURN	1116	1117
FT	STRAND	1120	1120
FT	STRAND	1122	1122
FT	TURN	1123	1124
FT	STRAND	1127	1131
FT	HELIX	1132	1134
FT	TURN	1135	1136
FT	STRAND	1137	1141
FT	TURN	1144	1145
FT	STRAND	1148	1152
FT	STRAND	1155	1158
FT	TURN	1159	1162

BY SIMILARITY.
 ZINC (CATALYTIC) (BY SIMILARITY).
 POTENTIAL.
 POTENTIAL.
 INTERCHAIN.

```

FT  STRAND      1163    1166
FT  STRAND      1173    1178
FT  TURN        1184    1185
FT  STRAND      1188    1188
FT  STRAND      1190    1190
FT  TURN        1191    1192
FT  STRAND      1193    1201

```

```

Query Match          100.0%;  Score 74;  DB 1;  Length 1314;
Best Local Similarity 100.0%;  Pred. No. 1.8e-05;
Matches   15;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

```

```

Qy          1 QYIKANSKFIGITEL 15
            |||||
Db          829 QYIKANSKFIGITEL 843

```

RESULT 2

VG84_BPML5

ID VG84_BPML5 STANDARD; PRT; 66 AA.

AC Q05301;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-FEB-1994 (Rel. 28, Last annotation update)

DE Gene 84 protein (GP84).

GN 84.

OS Mycobacteriophage L5.

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;

OC L5-like viruses.

OX NCBI_TaxID=31757;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93211282; PubMed=8459766;

RA Hatfull G.F., Sarkis G.J.;

RT "DNA sequence, structure and gene expression of mycobacteriophage L5:

RT a phage system for mycobacterial genetics.";

RL Mol. Microbiol. 7:395-405(1993).

CC

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Z18946; CAA79460.1; -.

DR PIR; S31029; S31029.

SQ SEQUENCE 66 AA; 7424 MW; 9C7104C7A4FA74A5 CRC64;

```

Query Match          59.5%;  Score 44;  DB 1;  Length 66;
Best Local Similarity 57.1%;  Pred. No. 0.29;
Matches    8;  Conservative   2;  Mismatches   4;  Indels   0;  Gaps   0;

```

```

Qy          2 YIKANSKFIGITEL 15
            ||| | ||:| |:
Db          50 YIKRNGKFGVTWEV 63

```

RESULT 3

MSP1_PLAF3

ID MSP1_PLAF3 STANDARD; PRT; 1682 AA.
AC P19598; Q25921;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (PMMSA) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate ro-33 / Ghana).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5834;
RN [1]
RP SEQUENCE OF 1-1061 FROM N.A.
RX MEDLINE=88166657; PubMed=3327688;
RA Certa U., Rotmann D., Matile H., Reber-Liske R.;
RT "A naturally occurring gene encoding the major surface antigen
RT precursor p190 of Plasmodium falciparum lacks tripeptide repeats.";
RL EMBO J. 6:4137-4142(1987).
RN [2]
RP SEQUENCE OF 1032-1682 FROM N.A.
RX MEDLINE=95354793; PubMed=7628566;
RA Tolle R., Bujard H., Cooper J.A.;
RT "Plasmodium falciparum: variations within the C-terminal region of
RT merozoite surface antigen-1.";
RL Exp. Parasitol. 81:47-54(1995).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
CC kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M35727; AAA29715.1; -.
DR EMBL; Y00087; CAA68280.1; -.
DR EMBL; Z35326; CAA84555.1; -.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1682 MEROZOITE SURFACE PROTEIN 1.
FT TRANSMEM 1666 1682 MEMBRANE ANCHOR (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 599 599 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	CARBOHYD	785	785	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	881	881	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	901	901	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	947	947	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1071	1071	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1178	1178	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1569	1569	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1682 AA; 192462 MW; C82A1E159948CAD6 CRC64;		

Query Match 57.4%; Score 42.5; DB 1; Length 1682;
 Best Local Similarity 60.0%; Pred. No. 13;
 Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

Qy 1 QYIKANSKFI-GITE 14
 |::|:|||| | |::||
 Db 983 QFVKSNSKVITGLTE 997

RESULT 4

MSP1_PLAFF

ID MSP1_PLAFF STANDARD; PRT; 1701 AA.
 AC P13819;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
 DE (PMMSA).
 GN MSP-1.
 OS Plasmodium falciparum (isolate FC27 / Papua New Guinea).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88142999; PubMed=2449612;
 RA Peterson M.G., Coppel R.L., McIntyre P., Langford C.J., Woodrow G.,
 RA Brown G.V., Anders R.F., Kemp D.J.;
 RT "Variation in the precursor to the major merozoite surface antigens
 RT of Plasmodium falciparum."
 RL Mol. Biochem. Parasitol. 27:291-302(1988).
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Potential).
 CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
 CC kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
 CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M19143; AAA29653.1; -.
 DR PIR; A54498; A54498.
 DR InterPro; IPR006209; EGF_like.
 DR Pfam; PF00008; EGF; 1.

KW	Malaria; Merozoite; Polypeptide; Repeat; Signal; Glycoprotein;			
KW	Transmembrane; GPI-anchor.			
FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	1701	MEROZOITE SURFACE PROTEIN 1.
FT	CARBOHYD	110	110	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	470	470	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	536	536	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	607	607	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	802	802	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	899	899	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	919	919	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	965	965	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	991	991	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1089	1089	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1196	1196	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1588	1588	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1701 AA; 193719 MW; 3920B75E73D38552 CRC64;		

```

RESULT 5
MSP1_PLAFM
ID   MSP1_PLAFM          STANDARD;          PRT;   1701 AA.
AC   P08569;
DT   01-AUG-1988 (Rel. 08, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE   (PMMSA) (P190).
GN   MSP-1.
OS   Plasmodium falciparum (isolate mad20 / Papua New Guinea).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=70153;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88011243; PubMed=3079521;
RA   Tanabe K., Mackay M., Goman M., Scaife J.G.;
RT   "Allelic dimorphism in a surface antigen gene of the malaria parasite
RT   Plasmodium falciparum.";
RL   J. Mol. Biol. 195:273-287(1987).
RN   [2]
RP   REVISIONS TO 1403; 1569 AND 1629.
RA   Tanabe K.;
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-115 FROM N.A.
RX   MEDLINE=86136024; PubMed=3004972;
RA   Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
RA   Stunnenberg H., Bujard H.;

```

RT "Polymorphism of the precursor for the major surface antigens of
RT Plasmodium falciparum merozoites: studies at the genetic level.";
RL EMBO J. 4:3823-3829(1985).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
CC kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05624; CAA29112.1; -.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1701 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 802 802 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 899 899 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 919 919 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 991 991 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1089 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1196 1196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1588 1588 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1701 AA; 193768 MW; 3FC2EC59AF96EA98 CRC64;

Query Match 57.4%; Score 42.5; DB 1; Length 1701;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 QYIKANSKFI-GITE 14
|::|:|||| | |::|
Db 1001 QFVKSNSKVITGLTE 1015

RESULT 6
MSP1_PLAFC
ID MSP1_PLAFC STANDARD; PRT; 1726 AA.
AC P04934;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (PMMSA) (P195).

GN MSP-1.
 OS Plasmodium falciparum (isolate Camp / Malaysia).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5835;
 RN [1]
 RP SEQUENCE OF 1-1103 FROM N.A.
 RX MEDLINE=86205236; PubMed=3517809;
 RA Weber J.L., Leininger W.M., Lyon J.A.;
 RT "Variation in the gene encoding a major merozoite surface antigen of
 RT the human malaria parasite Plasmodium falciparum.";
 RL Nucleic Acids Res. 14:3311-3323(1986).
 RN [2]
 RP SEQUENCE OF 1104-1726 FROM N.A.
 RX MEDLINE=88143999; PubMed=3278296;
 RA Weber J.L., Sim B.K.L., Lyon J.A., Wolff R.;
 RT "Merozoite surface protein sequence from the Camp strain of the human
 RT malaria parasite Plasmodium falciparum.";
 RL Nucleic Acids Res. 16:1206-1206(1988).
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Potential).
 CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
 CC kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
 CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X03831; CAA27446.1; -.
 DR PIR; A23386; SAZQGM.
 DR InterPro; IPR006209; EGF_like.
 DR Pfam; PF00008; EGF; 1.
 KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
 KW Transmembrane; GPI-anchor.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 1726 MEROZOITE SURFACE PROTEIN 1.
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 827 827 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 924 924 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 944 944 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 990 990 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1221 1221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1613 1613 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1726 AA; 196197 MW; DD8AD45FA352BCF3 CRC64;

Query Match 57.4%; Score 42.5; DB 1; Length 1726;
 Best Local Similarity 60.0%; Pred. No. 14;

Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 QYIKANSKFI-GITE 14
|::|:| | | | | | | | | |
Db 1026 QFVKSNSKVVITGLTE 1040

RESULT 7

MSP1_PLAFP

ID MSP1_PLAFP STANDARD; PRT; 1726 AA.
AC P50495;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (PMMSA) (GP195).
GN MSP-1.
OS Plasmodium falciparum (isolate Palo Alto / Uganda).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57270;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89005525; PubMed=3049134;
RA Chang S.P., Kramer K.J., Yamaga K.M., Kato A., Case S.E.,
RA Siddiqui W.A.;
RT "Plasmodium falciparum: gene structure and hydropathy profile of the
RT major merozoite surface antigen (gpl95) of the Uganda-Palo Alto
RT isolate.";
RL Exp. Parasitol. 67:1-11(1988).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
CC kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M37213; AAA29611.1; -.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1726 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 827 827 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 924 924 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 944 944 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 990 990 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1221 1221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1613 1613 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1726 AA; 196174 MW; 5B59CEEFA2F9A026 CRC64;

Query Match 57.4%; Score 42.5; DB 1; Length 1726;
 Best Local Similarity 60.0%; Pred. No. 14;
 Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 QYIKANSKFI-GITE 14
 |::|:|||| | |::|
 Db 1026 QFVKSNSKVITGLTE 1040

RESULT 8

PYRC_SERMA

ID PYRC_SERMA STANDARD; PRT; 204 AA.
 AC Q9S3S1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Dihydroorotase (EC 3.5.2.3) (DHOase) (Fragment).
 GN PYRC.
 OS *Serratia marcescens*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Serratia*.
 OX NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SM6;
 RA Berkmen M., Benedik M.J.;
 RT "DinI inhibits transcription of *Serratia marcescens* nuclease.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
 CC -!- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-aspartate.
 CC -!- COFACTOR: This enzyme tightly binds one zinc atom per chain which is required for the catalytic mechanism. It also binds weakly to two other zincs which are not essential for activity (By similarity).
 CC -!- PATHWAY: Pyrimidine biosynthesis; third step.
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE DHOASE FAMILY. SUBFAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF175466; AAD50307.1; ALT_INIT.
 DR HAMAP; MF_00219; -, 1.
 DR InterPro; IPR006680; Amidohydro_1.

DR InterPro; IPR002195; Dihydroorotase.
 DR Pfam; PF01979; Amidohydro_1; 1.
 DR PROSITE; PS00482; DIHYDROOROTASE_1; PARTIAL.
 DR PROSITE; PS00483; DIHYDROOROTASE_2; 1.
 KW Pyrimidine biosynthesis; Hydrolase; Zinc.
 FT NON_TER 1 1
 SQ SEQUENCE 204 AA; 22899 MW; 9DB0B6C9B834B310 CRC64;

Query Match 55.4%; Score 41; DB 1; Length 204;
 Best Local Similarity 46.2%; Pred. No. 3.1;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGIT 13
 ||::| ::|:| |
 Db 42 QYVQAGNRFLGAT 54

RESULT 9

LE12_THETN

ID LE12_THETN STANDARD; PRT; 384 AA.
 AC Q8RCF9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE 2-isopropylmalate synthase 2 (EC 2.3.3.13) (Alpha-isopropylmalate
 DE synthase 2) (Alpha-IPM synthetase 2).
 GN LEUA2 OR TTE0472.
 OS Thermoanaerobacter tengcongensis.
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.
 OX NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4 / JCM 11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of T. tengcongensis genome.";
 RL Genome Res. 12:689-700(2002).
 CC -!- FUNCTION: Catalyzes the condensation of the acetyl group of
 CC acetyl-CoA with 3-methyl-2-oxobutanoate (2-oxoisovalerate) to form
 CC 3-carboxy-3-hydroxy-4-methylpentanoate (2-isopropylmalate).
 CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + 3-methyl-2-oxobutanoate + H(2)O =
 CC 2-hydroxy-2-isopropylsuccinate + CoA.
 CC -!- PATHWAY: Leucine biosynthesis; first step.
 CC -!- SUBUNIT: Homotetramer (By similarity).
 CC -!- SIMILARITY: Belongs to the alpha-IPM synthetase / homocitrate
 CC synthase family. LeuA 1 subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL; AE013018; AAM23753.1; -.
DR HAMAP; MF_01025; atypical; 1.
DR InterPro; IPR002034; AIPM/Hcit_synth.
DR InterPro; IPR000891; HMGL-like.
DR Pfam; PF00682; HMGL-like; 1.
DR PROSITE; PS00815; AIPM_HOMOCIT_SYNTH_1; 1.
DR PROSITE; PS00816; AIPM_HOMOCIT_SYNTH_2; 1.
KW Leucine biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 384 AA; 42404 MW; 095310F2C0E4A4DD CRC64;

```

```

Query Match          55.4%; Score 41; DB 1; Length 384;
Best Local Similarity 54.5%; Pred. No. 5.8;
Matches      6; Conservative      4; Mismatches      1; Indels      0; Gaps      0;

```

```

Qy      3 IKANSKFIGIT 13
       :|| :||:|:|
Db      214 VKAGAKFVGVT 224

```

RESULT 10

V120_HSVSA

```

ID V120_HSVSA STANDARD; PRT; 899 AA.
AC Q01055;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Capsid assembly protein 63.
GN 63 OR EERF1.
OS Herpesvirus saimiri (strain 11).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10383;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92333688; PubMed=1321287;
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA Honess R.W.;
RT "Primary structure of the herpesvirus saimiri genome.";
RL J. Virol. 66:5047-5058(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92230228; PubMed=1314457;
RA Nicholas J., Cameron K.R., Coleman H., Newman C., Honess R.W.;
RT "Analysis of nucleotide sequence of the rightmost 43 kbp of
RT herpesvirus saimiri (HVS) L-DNA: general conservation of genetic
RT organization between HVS and Epstein-Barr virus.";
RL Virology 188:296-310(1992).
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL37,
CC EHV-1 23, EBV BOLF1, VZV 21, HVS-1 63, AND HCMV UL47.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X64346; CAA45686.1; -.
DR EMBL; M86409; AAA46139.1; -.
KW Capsid assembly.
SQ SEQUENCE 899 AA; 103350 MW; F1429B3770A2885E CRC64;

Query Match 55.4%; Score 41; DB 1; Length 899;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14
||| :|: | |:|
Db 124 QYITSNATFTGLSE 137

RESULT 11

ACPD_HAEIN

ID ACPD_HAEIN STANDARD; PRT; 194 AA.
AC P43013;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable acyl carrier protein phosphodiesterase (EC 3.1.4.14) (ACP
DE phosphodiesterase).
GN ACPD OR HI1366.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=96186898; PubMed=8635745;
RA Chandler M.S., Smith R.A.;
RT "Characterization of the Haemophilus influenzae topA locus: DNA
RT topoisomerase I is required for genetic competence.";
RL Gene 169:25-31(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of
CC the phosphopantetheine prosthetic group from ACP (By similarity).

CC -!- CATALYTIC ACTIVITY: Holo-[acyl-carrier protein] + H(2)O = 4'-
CC phosphopantetheine + apo-[acyl-carrier protein].
CC -!- SIMILARITY: Belongs to the acpD family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U20964; AAC43728.1; -.
DR EMBL; U32816; AAC23013.1; -.
DR PIR; G64026; G64026.
DR TIGR; HI1366; -.
DR HAMAP; MF_01216; -; 1.
DR InterPro; IPR003680; NADHdh_2.
DR Pfam; PF02525; Flavodoxin_2; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 194 AA; 21208 MW; A8AEC9D00829522C CRC64;

Query Match 54.1%; Score 40; DB 1; Length 194;
Best Local Similarity 53.3%; Pred. No. 4.5;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 ||:|: |||||::
Db 147 QYMKSilGFIGITDV 161

RESULT 12

PEF1_LACLC

ID PEF1_LACLC STANDARD; PRT; 601 AA.
AC P54124; P94880;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oligoendopeptidase F, plasmid (EC 3.4.24.-).
GN PEPF1 OR PEPF.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OG Plasmid pLP763.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1359;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=NCDO 763;
RX MEDLINE=95096044; PubMed=7798200;
RA Monnet V., Nardi M., Chopin A., Chopin M.-C., Gripon J.-C.;
RT "Biochemical and genetic characterization of PepF, an oligopeptidase
RT from Lactococcus lactis.";
RL J. Biol. Chem. 269:32070-32076(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NCDO 763;
RX MEDLINE=97352670; PubMed=9209029;
RA Nardi M., Renault P., Monnet V.;

RT "Duplication of the pepF gene and shuffling of DNA fragments on the
RT lactose plasmid of Lactococcus lactis.";
RL J. Bacteriol. 179:4164-4171(1997).
CC -!- FUNCTION: HYDROLYZES PEPTIDES CONTAINING BETWEEN 7 AND 17 AMINO
CC ACIDS WITH A RATHER WIDE SPECIFICITY.
CC -!- COFACTOR: Binds 1 zinc ion (By similarity).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z32522; CAA83534.1; -.
DR EMBL; X99798; CAA68133.1; -.
DR PIR; A55485; A55485.
DR MEROPS; M03.007; -.
DR InterPro; IPR004438; PepF.
DR InterPro; IPR001567; Peptidase_M3.
DR InterPro; IPR006025; Zn_MTpeptdse.
DR Pfam; PF01432; Peptidase_M3; 1.
DR TIGRFAMS; TIGR00181; pepF; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Plasmid.
FT METAL 387 387 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 388 388 BY SIMILARITY.
FT METAL 391 391 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 394 394 ZINC (CATALYTIC) (BY SIMILARITY).
FT CONFLICT 518 518 F -> S (IN REF. 2).
SQ SEQUENCE 601 AA; 69674 MW; C8B5C519FFA2F787 CRC64;

Query Match 54.1%; Score 40; DB 1; Length 601;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
:|: | :||:|
Db 284 RYIELRKKILGITDL 298

RESULT 13
PEPF_LACLA
ID PEPF_LACLA STANDARD; PRT; 601 AA.
AC Q9CEV7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oligoendopeptidase F homolog (EC 3.4.24.-).
GN PEPF OR LL1727.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium *Lactococcus*
RT *lactis* ssp. *lactis* IL1403.";
RL Genome Res. 11:731-753(2001).
CC -!- FUNCTION: HYDROLYZES PEPTIDES CONTAINING BETWEEN 7 AND 17 AMINO
CC ACIDS WITH A RATHER WIDE SPECIFICITY (BY SIMILARITY).
CC -!- COFACTOR: Binds 1 zinc ion (By similarity).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006403; AAK05825.1; -.
DR PIR; G86840; G86840.
DR InterPro; IPR004438; PepF.
DR InterPro; IPR001567; Peptidase_M3.
DR InterPro; IPR006025; Zn_MTpeptdse.
DR Pfam; PF01432; Peptidase_M3; 1.
DR TIGRFAMs; TIGR00181; pepF; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Complete proteome.
FT METAL 387 387 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 388 388 BY SIMILARITY.
FT METAL 391 391 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 394 394 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 601 AA; 69883 MW; 045E6873F10FE19A CRC64;

Query Match 54.1%; Score 40; DB 1; Length 601;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
:| |: | :||| :|
Db 284 RYIELRKKILGITDL 298

RESULT 14
YHJ9_YEAST
ID YHJ9_YEAST STANDARD; PRT; 644 AA.
AC P38694;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical aldehyde-dehydrogenase like protein in FIL1-VMA10
DE intergenic region (EC 1.2.1.-).
GN YHR039C.
OS *Saccharomyces cerevisiae* (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohldman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RT "Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome
 RT VIII.";
 RL Science 265:2077-2082(1994).
 CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U00062; AAB68915.1; -.
 DR PIR; S46746; S46746.
 DR SGD; S0001081; MSC7.
 DR GO; GO:0007131; P:meiotic recombination; IMP.
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR Pfam; PF00171; aldedh; 1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
 KW Hypothetical protein; Oxidoreductase.
 FT ACT_SITE 354 354 BY SIMILARITY.
 FT ACT_SITE 389 389 BY SIMILARITY.
 SQ SEQUENCE 644 AA; 71320 MW; 54DADDAEB2A16D4D CRC64;

Query Match 54.1%; Score 40; DB 1; Length 644;
 Best Local Similarity 60.0%; Pred. No. 15;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 | | : | | | | |
 Db 38 QIIQDNQKLIGITT 52

RESULT 15
 KAD_MYCCA
 ID KAD_MYCCA STANDARD; PRT; 213 AA.
 AC P10251;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Adenylate kinase (EC 2.7.4.3) (ATP-AMP transphosphorylase).
 GN ADK.
 OS *Mycoplasma capricolum*.

OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
 OC Entomoplasmataceae.
 OX NCBI_TaxID=2095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 27343 / Kid;
 RX MEDLINE=88142549; PubMed=3481422;
 RA Ohkubo S., Muto A., Kawauchi Y., Yamao F., Osawa S.;
 RT "The ribosomal protein gene cluster of Mycoplasma capricolum.";
 RL Mol. Gen. Genet. 210:314-322(1987).
 CC -!- FUNCTION: THIS SMALL UBIQUITOUS ENZYME IS ESSENTIAL FOR
 CC MAINTENANCE AND CELL GROWTH.
 CC -!- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the adenylate kinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X06414; CAA29724.1; -.
 DR PIR; S02851; KIYMC.
 DR HSSP; P27142; 1ZIN.
 DR HAMAP; MF_00235; -; 1.
 DR InterPro; IPR006259; Adenyl_kin_sub.
 DR InterPro; IPR000850; Adenylate_kin.
 DR Pfam; PF00406; ADK; 1.
 DR Pfam; PF05191; ADK_lid; 1.
 DR PRINTS; PR00094; ADENYLTKNASE.
 DR ProDom; PD000657; Adenylate_kin; 1.
 DR TIGRFAMS; TIGR01351; adk; 1.
 DR PROSITE; PS00113; ADENYLATE_KINASE; 1.
 KW Transferase; Kinase; ATP-binding.
 FT NP_BIND 7 15 ATP (BY SIMILARITY).
 SQ SEQUENCE 213 AA; 24616 MW; FBCFA426B6F92E16 CRC64;

Query Match 52.7%; Score 39; DB 1; Length 213;
 Best Local Similarity 72.7%; Pred. No. 7.6;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 YIKANSKFIGI 12
 | | | | | | |
 Db 183 YFKTNSKFIEI 193

Search completed: December 13, 2003, 13:35:06
 Job time : 2.51282 secs

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 0.934066 Seconds
(without alignments)
1544.356 Million cell updates/sec

Title: US-09-785-215-4
Perfect score: 74
Sequence: 1 QYIKANSKFIGITEL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query Match	Length			
1	74	100.0	1315	1	BTCLTN	tentoxilysin (EC 3
2	44.5	60.1	244	2	S29982	class II histocomp
3	44	59.5	66	2	S31029	gene 84 protein -
4	43	58.1	180	2	G86826	diamine N-acetyltr
5	43	58.1	899	2	T42976	hypothetical prote
6	42.5	57.4	1060	2	S06286	major merozoite su
7	42.5	57.4	1086	2	S16752	major merozoite su
8	42.5	57.4	1701	2	A54498	major merozoite su
9	42.5	57.4	1701	2	A26868	major merozoite su
10	42.5	57.4	1726	1	SAZQGM	major merozoite su
11	42.5	57.4	1726	2	A45948	major merozoite su
12	42	56.8	1333	2	S38635	blastopia polyprot
13	41	55.4	123	2	G48677	Ig heavy chain V-D

14	41	55.4	447	2	H97146	siderophore/Surfac
15	41	55.4	899	2	G36812	hypothetical prote
16	40.5	54.7	245	2	S29980	class II histocomp
17	40	54.1	79	2	D85794	hypothetical prote
18	40	54.1	194	2	G64026	acyl-carrier-prot
19	40	54.1	601	1	A55485	oligopeptidase (EC
20	40	54.1	601	2	G86840	oligoendopeptidase
21	40	54.1	644	2	S46746	hypothetical prote
22	39	52.7	102	2	PH1491	Ig heavy chain V r
23	39	52.7	119	2	PH1516	Ig heavy chain V r
24	39	52.7	119	2	PH1518	Ig heavy chain V r
25	39	52.7	119	2	PH1519	Ig heavy chain V r
26	39	52.7	123	2	F48677	Ig heavy chain V-D
27	39	52.7	135	2	PH1494	Ig heavy chain V r
28	39	52.7	140	2	PH1488	Ig heavy chain V r
29	39	52.7	189	2	G97978	conserved hypothet
30	39	52.7	213	1	KIYMC	adenylate kinase (
31	39	52.7	326	2	B71808	type II restrictio
32	39	52.7	349	2	T43043	probable acetyl-Co
33	39	52.7	423	2	F64690	type IIS restricti
34	39	52.7	505	2	C90569	hypothetical prote
35	38	51.4	188	2	H71875	hypothetical prote
36	38	51.4	188	2	A64639	hypothetical prote
37	38	51.4	256	2	F64472	hypothetical prote
38	38	51.4	287	2	F70361	tRNA-pseudouridine
39	38	51.4	381	2	F71196	probable hexosyltr
40	38	51.4	383	2	T51466	hypothetical prote
41	38	51.4	424	2	T29127	hypothetical prote
42	38	51.4	501	2	A86158	endo-1,4-beta gluc
43	38	51.4	501	2	T52135	cellulase (EC 3.2.
44	38	51.4	561	2	E82395	methyl-accepting c
45	38	51.4	572	1	HNNZ80	hemagglutinin-neur

ALIGNMENTS

RESULT 1

BTCLTN

tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani

N;Alternate names: tetanus neurotoxin

C;Species: Clostridium tetani

C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 03-Jun-2002

C;Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364

R;Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, M.; Habermann, E.; Niemann, H.

EMBO J. 5, 2495-2502, 1986

A;Title: Tetanus toxin: primary structure, expression in E. coli, and homology with botulinum toxins.

A;Reference number: A25689; MUID:87053814; PMID:3536478

A;Accession: A25689

A;Molecule type: DNA

A;Residues: 1-1315 <EIS>

A;Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770

R;Fairweather, N.F.; Lyness, V.A.

Nucleic Acids Res. 14, 7809-7812, 1986

A;Title: The complete nucleotide sequence of tetanus toxin.

A;Reference number: A25757; MUID:87040747; PMID:3774547
 A;Accession: A25757
 A;Molecule type: DNA
 A;Residues: 1-1315 <FAI>
 A;Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774
 A;Experimental source: strain CN3911
 R;Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
 J. Bacteriol. 165, 21-27, 1986
 A;Title: Cloning, nucleotide sequencing, and expression of tetanus toxin
 fragment C in Escherichia coli.
 A;Reference number: A25194; MUID:86085672; PMID:3510187
 A;Accession: A25194
 A;Molecule type: DNA
 A;Residues: 743-1315 <FA2>
 A;Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921
 A;Accession: B25194
 A;Molecule type: protein
 A;Residues: 865-894 <FA3>
 R;Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
 Infect. Immun. 57, 3588-3593, 1989
 A;Title: Isolation, purification, and characterization of fragment B, the NH-2-
 terminal half of the heavy chain of tetanus toxin.
 A;Reference number: A60759; MUID:90035436; PMID:2478476
 A;Accession: A60759
 A;Molecule type: protein
 A;Residues: 461-475 <MAT>
 R;Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin,
 G.
 J. Immunol. 142, 394-402, 1989
 A;Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
 A;Reference number: JS0098; MUID:89093918; PMID:2463305
 A;Contents: annotation; epitope region
 R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.;
 DasGupta, B.R.; Montecucco, C.
 Nature 359, 832-835, 1992
 A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by
 proteolytic cleavage of synaptobrevin.
 A;Reference number: S27125; MUID:93063293; PMID:1331807
 A;Contents: annotation
 R;de Philippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
 Eur. J. Biochem. 229, 61-69, 1995
 A;Title: Structural studies on the zinc-endopeptidase light chain of tetanus
 neurotoxin.
 A;Reference number: S69348; MUID:95262688; PMID:7744050
 A;Accession: S69348
 A;Molecule type: protein
 A;Residues: 2-31 <DEF>
 C;Comment: The source of this protein was an extrachromosomal plasmid.
 C;Comment: The precursor is cleaved by endogenous proteinase activity to form
 light (fragment A) and heavy (fragment B.C) chains that are covalently linked by
 an interchain disulfide bond (the individual chains are not toxic when
 separated). The amino end of the heavy chain (fragment B) can be separated from
 the carboxyl end (fragment C) by papain.
 C;Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to
 gangliosides and may target the toxin to the motor end plate. Fragment A is a
 zinc-dependent endopeptidase.

C;Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized, and moves by retrograde transport up the axon into the spinal cord where it can move between postsynaptic and presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of synaptobrevin (vesicle-associated membrane protein 2).

C;Function:

A;Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2

C;Superfamily: tetanus toxin

C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc

F;2-457/Product: tentoxilysin light chain (fragment A) #status predicted <TTL>

F;461-1315/Product: tentoxilysin heavy chain (fragment B.C) #status experimental <TTH>

F;461-864/Domain: channel forming (fragment B) #status predicted <TXB>

F;865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>

F;233,237/Binding site: zinc (His) #status predicted

F;234/Active site: Glu #status predicted

Query Match 100.0%; Score 74; DB 1; Length 1315;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | | |
Db 830 QYIKANSKFIGITEL 844

RESULT 2

S29982

class II histocompatibility antigen - Atlantic salmon

C;Species: Salmo salar (Atlantic salmon)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999

C;Accession: S29982

R;Hordvik, I.

submitted to the EMBL Data Library, October 1992

A;Reference number: S29980

A;Accession: S29982

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-244 <HOR>

A;Cross-references: EMBL:X70166; NID:g64369; PID:g64370

C;Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 60.1%; Score 44.5; DB 2; Length 244;
Best Local Similarity 55.6%; Pred. No. 1.8;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 1 QYIKANS---KFIGITEL 15
: || : || : || : ||
Db 51 EYIRFNSTVGKFGVGYTEL 68

RESULT 3

S31029

gene 84 protein - Mycobacterium phage L5

C;Species: Mycobacterium phage L5

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999
 C;Accession: S31029
 R;Donnelly-Wu, M.K.; Jacobs Jr., W.R.; Hatfull, G.F.
 Mol. Microbiol. 7, 407-417, 1993
 A;Title: Superinfection immunity of mycobacteriophage L5: applications for genetic transformation of mycobacteria.
 A;Reference number: S30949; MUID:93211283; PMID:8459767
 A;Accession: S31029
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-66 <DON>
 A;Cross-references: EMBL:Z18946; NID:g15859; PIDN:CAA79460.1; PID:e59702; PID:g579152
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1992
 C;Genetics:
 A;Gene: 84
 A;Start codon: GTG

Query Match 59.5%; Score 44; DB 2; Length 66;
 Best Local Similarity 57.1%; Pred. No. 0.59;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 YIKANSKFIGITEL 15
 ||| | ||:| |:
 Db 50 YIKRNGKFGVTWEV 63

RESULT 4

G86826

diamine N-acetyltransferase (EC 2.3.1.57) [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

N;Alternate names: spermidine acetyltransferase

C;Species: Lactococcus lactis subsp. lactis

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 17-May-2002

C;Accession: G86826

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich, S.D.; Sorokin, A.

Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis IL1403.

A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: G86826

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-180 <STO>

A;Cross-references: GB:AE005176; PID:g12724622; PIDN:AAK05713.1; GSPDB:GN00146

A;Experimental source: strain IL1403

C;Genetics:

A;Gene: yqfF

C;Superfamily: Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimJ

C;Keywords: acyltransferase; coenzyme A

Query Match 58.1%; Score 43; DB 2; Length 180;
 Best Local Similarity 69.2%; Pred. No. 2.5;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 IKANSKFIGITEL 15
|:| | ||| |
Db 65 IEANDTFIGIVEL 77

RESULT 5

T42976

hypothetical protein 63 - ateline herpesvirus 3 (strain 73)

C;Species: ateline herpesvirus 3

A;Variety: strain 73

C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000

C;Accession: T42976

R;Albrecht, J.C.; Fleckenstein, B.

submitted to the EMBL Data Library, August 1998

A;Description: Primary structure of the herpesvirus ateles genome.

A;Reference number: Z22274

A;Accession: T42976

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-899 <ALB>

A;Cross-references: EMBL:AF083424; PIDN:AAC95587.1

A;Experimental source: strain 73

Query Match 58.1%; Score 43; DB 2; Length 899;
Best Local Similarity 64.3%; Pred. No. 13;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14
||| :|| | | |
Db 124 QYITSNSTFTGQTE 137

RESULT 6

S06286

major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (strain RO-33 Ghana) (fragment)

N;Alternate names: 190K protein; polymorphic schizont antigen

C;Species: Plasmodium falciparum

C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jun-2000

C;Accession: S06286

R;Certa, U.; Rotmann, D.; Matile, H.; Reber-Liske, R.

EMBO J. 6, 4137-4142, 1987

A;Title: A naturally occurring gene encoding the major surface antigen precursor p190 of Plasmodium falciparum lacks tripeptide repeats.

A;Reference number: S06286; MUID:88166657; PMID:3327688

A;Accession: S06286

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-1060 <CER>

C;Superfamily: major merozoite surface antigen

C;Keywords: surface antigen

Query Match 57.4%; Score 42.5; DB 2; Length 1060;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

Qy 1 QYIKANSKFI-GITE 14
|::|:|||| | |::||
Db 983 QFVKSNSKVVITGLTE 997

RESULT 7

S16752

major merozoite surface antigen MSA-1 - malaria parasite (*Plasmodium falciparum*) (fragments)

N;Alternate names: polymorphic schizont antigen p190

C;Species: *Plasmodium falciparum*

C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jun-2000

C;Accession: A44865; S16752

R;Olafsson, P.; Matile, H.; Certa, U.

Exp. Parasitol. 74, 381-389, 1992

A;Title: *Plasmodium falciparum*: the repetitive MSA-1 surface protein of the RO-71 isolate is recognized by mouse antibody against the nonrepetitive repeat block of RO-33.

A;Reference number: A44865; MUID:92275047; PMID:1592091

A;Accession: A44865

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1086 <OLA>

A;Cross-references: EMBL:X61930

A;Experimental source: isolate RO-71

C;Genetics:

A;Gene: MSA1

C;Superfamily: major merozoite surface antigen

C;Keywords: glycoprotein; merozoite; surface antigen; tandem repeat

Query Match 57.4%; Score 42.5; DB 2; Length 1086;

Best Local Similarity 60.0%; Pred. No. 19;

Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

Qy 1 QYIKANSKFI-GITE 14
|::|:|||| | |::||
Db 1009 QFVKSNSKVVITGLTE 1023

RESULT 8

A54498

major merozoite surface antigen precursor - malaria parasite (*Plasmodium falciparum*) (isolate FC27)

C;Species: *Plasmodium falciparum*

C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000

C;Accession: A54498

R;Peterson, M.G.; Coppel, R.L.; McIntyre, P.; Langford, C.J.; Woodrow, G.; Brown, G.V.; Anders, R.F.; Kemp, D.J.

Mol. Biochem. Parasitol. 27, 291-302, 1988

A;Title: Variation in the precursor to the major merozoite surface antigens of *Plasmodium falciparum*.

A;Reference number: A54498; MUID:88142999; PMID:2449612

A;Accession: A54498

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1701 <PET>

A;Cross-references: GB:M19143; NID:g160412; PIDN:AAA29653.1; PID:g160413

C;Superfamily: major merozoite surface antigen
C;Keywords: surface antigen

Query Match 57.4%; Score 42.5; DB 2; Length 1701;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 QYIKANSKFI-GITE 14
|::|:|||| | |::||
Db 1001 QFVKSNSKVITGLTE 1015

RESULT 9

A26868

major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (strain MAD20/papua New Guinea)

C;Species: Plasmodium falciparum

C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jun-2000

C;Accession: A26868

R;Tanabe, K.; Mackay, M.; Goman, M.; Scaife, J.G.

J. Mol. Biol. 195, 273-287, 1987

A;Title: Allelic dimorphism in a surface antigen gene of the malaria parasite Plasmodium falciparum.

A;Reference number: A26868; MUID:88011243; PMID:3079521

A;Accession: A26868

A;Molecule type: DNA

A;Residues: 1-1701 <TAN>

C;Superfamily: major merozoite surface antigen

C;Keywords: surface antigen

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-1701/Product: major merozoite surface antigen #status predicted <MAT>

Query Match 57.4%; Score 42.5; DB 2; Length 1701;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 QYIKANSKFI-GITE 14
|::|:|||| | |::||
Db 1001 QFVKSNSKVITGLTE 1015

RESULT 10

SAZQGM

major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (strain CAMP/Malaysia)

N;Alternate names: 195K glycoprotein

C;Species: Plasmodium falciparum

C;Date: 30-Sep-1987 #sequence_revision 31-Mar-1991 #text_change 09-Jun-2000

C;Accession: A23386; S06361

R;Weber, J.L.; Leininger, W.M.; Lyon, J.A.

Nucleic Acids Res. 14, 3311-3323, 1986

A;Title: Variation in the gene encoding a major merozoite surface antigen of the human malaria parasite Plasmodium falciparum.

A;Reference number: A23386; MUID:86205236; PMID:3517809

A;Accession: A23386

A;Molecule type: DNA

A;Residues: 1-1104 <WEB1>

A;Cross-references: EMBL:X03831
 R;Weber, J.L.; Sim, B.K.L.; Lyon, J.A.; Wolff, R.
 Nucleic Acids Res. 16, 1206, 1988
 A;Title: Merozoite surface protein sequence from the Camp strain of the human malaria parasite Plasmodium falciparum.
 A;Reference number: S06361; MUID:88143999; PMID:3278296
 A;Accession: S06361
 A;Molecule type: DNA
 A;Residues: 1104-1726 <WEB2>
 A;Cross-references: EMBL:X03831
 C;Comment: The merozoite stages of different strains have strain-specific surface antigens that are involved in strain-specific immunity.
 C;Comment: P. falciparum has three stages: sporozoite, merozoite, and gametocyte. The merozoite stage can be recovered from the blood.
 C;Superfamily: major merozoite surface antigen
 C;Keywords: glycoprotein; merozoite; surface antigen; tandem repeat
 F;1-19/Domain: signal sequence #status predicted <SIG>
 F;20-1726/Product: major merozoite surface antigen #status predicted <MAT>
 F;67-87,91-96,100-105,109-120/Region: 3-residue repeats (S-G-T)
 F;757-765/Region: 3-residue repeats (T-E-E)
 F;133,272,501,567,638,827,839,924,944,990,1016,1114,1221,1613,1658/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 57.4%; Score 42.5; DB 1; Length 1726;
 Best Local Similarity 60.0%; Pred. No. 31;
 Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

Qy 1 QYIKANSKFI-GITE 14
 |::|:|||| | |::||
 Db 1026 QFVKSNSKVITGLTE 1040

RESULT 11

A45948
 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (strain Uganda-Palo Alto)
 C;Species: Plasmodium falciparum
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jun-2000
 C;Accession: A45948
 R;Chang, S.P.; Kramer, K.J.; Yamaga, K.M.; Kato, A.; Case, S.E.; Siddiqui, W.A.
 Exp. Parasitol. 67, 1-11, 1988
 A;Title: Plasmodium falciparum: gene structure and hydropathy profile of the major merozoite surface antigen (gp195) of the Uganda-Palo Alto isolate.
 A;Reference number: A45948; MUID:89005525; PMID:3049134
 A;Accession: A45948
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1726 <CHA>
 A;Cross-references: GB:M37213
 C;Superfamily: major merozoite surface antigen
 C;Keywords: surface antigen

Query Match 57.4%; Score 42.5; DB 2; Length 1726;
 Best Local Similarity 60.0%; Pred. No. 31;
 Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

Qy 1 QYIKANSKFI-GITE 14

Db |::|:|||| | |::||
1026 QFVKSNSKVVITGLTE 1040

RESULT 12

S38635

blastopia polyprotein - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000

C;Accession: S38635

R;Frommer, G.; Schuh, R.; Jdckle, H.

submitted to the EMBL Data Library, November 1993

A;Description: Localized expression of a novel micropia-like element in the blastoderm of *Drosophila melanogaster* is dependent on the anterior morphogen bicoid.

A;Reference number: S38635

A;Accession: S38635

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1333 <FRO>

A;Cross-references: EMBL:Z27119; NID:g415797; PID:g415798

C;Genetics:

A;Gene: FlyBase:micropia

A;Cross-references: FlyBase:FBgn0014947

C;Keywords: polyprotein

Query Match 56.8%; Score 42; DB 2; Length 1333;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
 |::|:| || || ||
Db 127 KYVQARSKMIGSAEL 141

RESULT 13

G48677

Ig heavy chain V-D-J region (419.1) - mouse (fragment)

C;Species: *Mus musculus* (house mouse)

C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999

C;Accession: G48677

R;Tassignon, J.; Brait, M.; Jamila, I.; Urbain, J.; Gottlieb, P.; Brown, A.; Hasemann, C.A.; Capra, J.D.; Meek, K.

Proc. Natl. Acad. Sci. U.S.A. 90, 9508-9512, 1993

A;Title: Molecular characterization of monoclonal CRI-A-positive anti-arsonate antibodies derived from idiotype-negative mice bearing a light chain polymorphism.

A;Reference number: A48677; MUID:94022404; PMID:8415731

A;Accession: G48677

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-123 <TAS>

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 55.4%; Score 41; DB 2; Length 123;

Best Local Similarity 64.3%; Pred. No. 4;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YIKANSKFIGITEL 15
||| | || | |
Db 57 YIKYNEKFKGTTTL 70

RESULT 14

H97146

siderophore/Surfactin synthetase related protein [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C;Accession: H97146

R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, H.M.; Dubois, J.; Qiu, D.; Hitti, J.; Wolf, Y.I.; Tatusov, R.L.; Sabathe, F.; Doucette-Stamm, L.; Soucaille, P.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum.

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: H97146

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-447 <KUR>

A;Cross-references: GB:AE001437; PIDN:AAK79963.1; PID:g15024986; GSPDB:GN00168

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC2004

Query Match 55.4%; Score 41; DB 2; Length 447;
Best Local Similarity 63.6%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIG 11
: ||: | |||
Db 291 KYIRTNKKFIG 301

RESULT 15

G36812

hypothetical protein ORF63 - saimiriine herpesvirus 1 (strain 11)

C;Species: saimiriine herpesvirus 1

A;Note: host Saimiri sciureus (common squirrel monkey)

C;Date: 16-Oct-1992 #sequence_revision 16-Oct-1992 #text_change 08-Oct-1999

C;Accession: G36812

R;Albrecht, J.

submitted to the EMBL Data Library, January 1992

A;Description: Primary structure of the herpesvirus saimiri genome.

A;Reference number: A36806

A;Accession: G36812

A;Molecule type: DNA

A;Residues: 1-899 <ALB>

A;Cross-references: GB:X64346; NID:g60320; PIDN:CAA45686.1; PID:g60384

R;Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.; Wittmann, S.; Craxton, M.A.; Coleman, H.; Fleckenstein, B.; Honess, R.W.

J. Virol. 66, 5047-5058, 1992

A;Title: Primary structure of the herpesvirus saimiri genome.

A;Reference number: A37309; MUID:92333688; PMID:1321287

A;Contents: annotation; protein-coding frames

A;Note: neither protein nor nucleotide sequence is given

C;Genetics:

A;Gene: 63

Query Match 55.4%; Score 41; DB 2; Length 899;

Best Local Similarity 50.0%; Pred. NO. 30;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14

||| :|: | |::|

Db 124 QYITSNATFTGLSE 137

Search completed: December 13, 2003, 13:30:53

Job time : 2.93407 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:30:00 ; Search time 1.7033 Seconds
 (without alignments)
 1637.856 Million cell updates/sec

Title: US-09-785-215-4
 Perfect score: 74
 Sequence: 1 QYIKANSKFIGITEL 15

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query					
No.	Score	Match Length DB ID					Description

1	74	100.0	15	9	US-09-862-849-2	Sequence 2, Appli
2	74	100.0	15	10	US-09-785-215-4	Sequence 4, Appli
3	74	100.0	15	12	US-10-237-656-13	Sequence 13, Appl
4	74	100.0	15	12	US-09-405-986-1	Sequence 1, Appli
5	74	100.0	15	12	US-10-223-809A-4	Sequence 4, Appli
6	74	100.0	15	12	US-10-261-446-19	Sequence 19, Appl
7	74	100.0	15	12	US-10-239-313A-618	Sequence 618, App
8	74	100.0	15	12	US-10-295-074-3	Sequence 3, Appli
9	74	100.0	15	15	US-10-204-362-4	Sequence 4, Appli
10	74	100.0	15	15	US-10-223-711-7	Sequence 7, Appli
11	74	100.0	16	9	US-09-848-834A-2	Sequence 2, Appli
12	74	100.0	16	12	US-09-930-915A-64	Sequence 64, Appl
13	74	100.0	16	12	US-10-082-014-145	Sequence 145, App
14	74	100.0	16	12	US-10-372-076-175	Sequence 175, App
15	74	100.0	17	11	US-09-865-294-3	Sequence 3, Appli
16	74	100.0	17	12	US-10-239-313A-619	Sequence 619, App
17	74	100.0	17	12	US-10-346-563-23	Sequence 23, Appl
18	74	100.0	19	12	US-10-239-313A-620	Sequence 620, App
19	74	100.0	27	12	US-10-076-674-7	Sequence 7, Appli
20	74	100.0	28	9	US-09-848-834A-11	Sequence 11, Appl
21	74	100.0	31	9	US-09-848-834A-15	Sequence 15, Appl
22	74	100.0	31	12	US-10-237-656-1	Sequence 1, Appli
23	74	100.0	32	12	US-10-237-656-5	Sequence 5, Appli
24	74	100.0	46	9	US-09-848-834A-19	Sequence 19, Appl
25	74	100.0	50	9	US-09-943-548-8	Sequence 8, Appli
26	74	100.0	50	15	US-10-339-522-8	Sequence 8, Appli
27	74	100.0	137	12	US-10-237-656-3	Sequence 3, Appli
28	74	100.0	158	12	US-10-297-942-4	Sequence 4, Appli
29	74	100.0	158	12	US-10-297-942-6	Sequence 6, Appli
30	74	100.0	158	12	US-10-297-942-8	Sequence 8, Appli
31	74	100.0	158	12	US-10-297-942-14	Sequence 14, Appl
32	74	100.0	158	12	US-10-297-942-16	Sequence 16, Appl
33	74	100.0	194	12	US-10-295-074-46	Sequence 46, Appl
34	74	100.0	194	12	US-10-295-074-47	Sequence 47, Appl
35	74	100.0	285	12	US-10-295-074-9	Sequence 9, Appli
36	74	100.0	285	12	US-10-295-074-11	Sequence 11, Appl
37	74	100.0	287	12	US-10-295-074-13	Sequence 13, Appl
38	74	100.0	287	12	US-10-295-074-15	Sequence 15, Appl
39	74	100.0	514	12	US-10-295-074-49	Sequence 49, Appl
40	74	100.0	514	12	US-10-295-074-51	Sequence 51, Appl
41	74	100.0	514	12	US-10-295-074-59	Sequence 59, Appl
42	74	100.0	517	12	US-10-295-074-53	Sequence 53, Appl
43	74	100.0	872	12	US-10-241-596-145	Sequence 145, App
44	74	100.0	879	12	US-10-241-596-143	Sequence 143, App
45	74	100.0	887	12	US-10-241-596-147	Sequence 147, App

ALIGNMENTS

RESULT 1
 US-09-862-849-2
 ; Sequence 2, Application US/09862849
 ; Patent No. US20020013274A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sudhir Paul

; APPLICANT: Larry J. Smith
; APPLICANT: Gennady Gololobov
; TITLE OF INVENTION: Methods for Identifying Inducers and Inhibitors of
Proteolytic
; TITLE OF INVENTION: Antibodies, Compositions and Their Uses
; FILE REFERENCE: UNMC 63123 DIV
; CURRENT APPLICATION NUMBER: US/09/862,849
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: US 09/046,373
; PRIOR FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-862-849-2

Query Match 100.0%; Score 74; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 2

US-09-785-215-4
; Sequence 4, Application US/09785215
; Publication No. US20020187157A1
; GENERAL INFORMATION:
; APPLICANT: JENSEN, Martin Roland et al.
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID
; FILE REFERENCE: 3631-0107P
; CURRENT APPLICATION NUMBER: US/09/785,215
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-785-215-4

Query Match 100.0%; Score 74; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 3

US-10-237-656-13
; Sequence 13, Application US/10237656


```
; Publication No. US20030152581A1
; GENERAL INFORMATION:
; APPLICANT: SAINT-REMY, Jean-Marie
; APPLICANT: JACQUEMIN, Marc
; TITLE OF INVENTION: COMPOUND AND METHOD FOR THE PREVENTION AND/OR THE
TREATMENT OF ALLERGY
; FILE REFERENCE: 2002-0771/LC/01699
; CURRENT APPLICATION NUMBER: US/10/237,656
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 09/362,731
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 13
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
US-10-237-656-13
```

```
Query Match          100.0%; Score 74; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 QYIKANSKFIGITEL 15
         |||||
Db      1 QYIKANSKFIGITEL 15
```

RESULT 4

```
US-09-405-986-1
; Sequence 1, Application US/09405986
; Publication No. US20030157115A1
; GENERAL INFORMATION:
; APPLICANT: BAY, Sylvie
; APPLICANT: CANTACUZENE, Daniele
; APPLICANT: LECLERC, Claude
; APPLICANT: LO-MAN, Richard
; TITLE OF INVENTION: MULTIPLE ANTIGEN GLYCOPEPTIDE CARBOHYDRATE, VACCINE
; TITLE OF INVENTION: COMPRISING THE SAME AND USE THEREOF
; FILE REFERENCE: 134I US 3565
; CURRENT APPLICATION NUMBER: US/09/405,986
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: 60/041,726
; EARLIER FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-405-986-1
```

```
Query Match          100.0%; Score 74; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 5

US-10-223-809A-4

; Sequence 4, Application US/10223809A
; Publication No. US20030157117A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; APPLICANT: Rasmussen, Peter Birk et al.
; TITLE OF INVENTION: No. US20030157117A1 Method for Down-Regulation of
Amyloid
; FILE REFERENCE: 674542-2008
; CURRENT APPLICATION NUMBER: US/10/223,809A
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: US 60/337,543
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: US 60/373,027
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: DE 2001 01231
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: DE 2002 0058
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-223-809A-4

Query Match 100.0%; Score 74; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 6

US-10-261-446-19

; Sequence 19, Application US/10261446
; Publication No. US20030165526A1
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE
THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261,446

; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-261-446-19

Query Match 100.0%; Score 74; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
|||
Db 1 QYIKANSKFIGITEL 15

RESULT 7

US-10-239-313A-618

; Sequence 618, Application US/10239313A
; Publication No. US20030175285A1
; GENERAL INFORMATION:
; APPLICANT: KLINGUER - HAMOUR, Christine
; APPLICANT: CORVAIA, Nathalie
; APPLICANT: BECK, Alain
; APPLICANT: GOETSCH, Liliane
; TITLE OF INVENTION: MOLECULE OF PHARMACEUTICAL INTEREST COMPRISING AT ITS
; TITLE OF INVENTION: N-TERMINAL A GLUTAMIC ACID OR A GLUTAMINE IN THE FORM
; TITLE OF INVENTION: OF A PHYSIOLOGICALLY ACCEPTABLE STRONG ACID
; FILE REFERENCE: 343 727 - US
; CURRENT APPLICATION NUMBER: US/10/239,313A
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: FR 00/03711
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT 01/70772
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 697
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 618
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-239-313A-618

Query Match 100.0%; Score 74; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
|||
Db 1 QYIKANSKFIGITEL 15

RESULT 8

US-10-295-074-3

; Sequence 3, Application US/10295074
 ; Publication No. US20030185845A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pharmexa A/S
 ; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
 ; FILE REFERENCE: P1013DK00
 ; CURRENT APPLICATION NUMBER: US/10/295,074
 ; CURRENT FILING DATE: 2002-11-15
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 15
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Tetanus toxoid P2 epitope
 US-10-295-074-3

Query Match 100.0%; Score 74; DB 12; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.4e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||
 Db 1 QYIKANSKFIGITEL 15

RESULT 9

US-10-204-362-4

; Sequence 4, Application US/10204362
 ; Publication No. US20030086938A1
 ; GENERAL INFORMATION:
 ; APPLICANT: M&E Biotech A/S
 ; TITLE OF INVENTION: No. US20030086938A1el Method For Down-Regulation Of Amyloid
 ; FILE REFERENCE: 3631-0120P
 ; CURRENT APPLICATION NUMBER: US/10/204,362
 ; CURRENT FILING DATE: 2002-08-16
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 3.0
 ; SEQ ID NO 4
 ; LENGTH: 15
 ; TYPE: PRT
 ; ORGANISM: Clostridium tetani
 US-10-204-362-4

Query Match 100.0%; Score 74; DB 15; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.4e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||
 Db 1 QYIKANSKFIGITEL 15

RESULT 10

US-10-223-711-7

; Sequence 7, Application US/10223711
 ; Publication No. US20030113344A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bakaletz, Lauren O.
 ; APPLICANT: Kaumaya, Pravin T.P.
 ; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
 ; FILE REFERENCE: 18525/04058
 ; CURRENT APPLICATION NUMBER: US/10/223,711
 ; CURRENT FILING DATE: 2002-08-19
 ; PRIOR APPLICATION NUMBER: 09/148,711
 ; PRIOR FILING DATE: 1998-09-04
 ; PRIOR APPLICATION NUMBER: 08/460,502
 ; PRIOR FILING DATE: 1995-06-02
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 15
 ; TYPE: PRT
 ; ORGANISM: Clostridium tetani
 US-10-223-711-7

Query Match 100.0%; Score 74; DB 15; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.4e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||
 Db 1 QYIKANSKFIGITEL 15

RESULT 11

US-09-848-834A-2

; Sequence 2, Application US/09848834A
 ; Patent No. US20020076416A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aphton Corporation
 ; TITLE OF INVENTION: Chimeric Peptide Immunogens
 ; FILE REFERENCE: 1102865-0047
 ; CURRENT APPLICATION NUMBER: US/09/848,834A
 ; CURRENT FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: 60/202,328
 ; PRIOR FILING DATE: 2000-05-05
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 16
 ; TYPE: PRT
 ; ORGANISM: Tetanus bacillus
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (1)..(16)
 ; OTHER INFORMATION: Amino acid sequence 829-844 of the Tetanus
 ; OTHER INFORMATION: Toxoid Precursor (Tentoxylisin)
 US-09-848-834A-2

Query Match 100.0%; Score 74; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | | | |
Db 2 QYIKANSKFIGITEL 16

RESULT 12

US-09-930-915A-64
; Sequence 64, Application US/09930915A
; Publication No. US20030138769A1
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: IMMUNOGENIC Hbc CHIMER PARTICLES HAVING ENHANCED
; TITLE OF INVENTION: STABILITY
; FILE REFERENCE: 4564/83501 ICC-102.2 PCT
; CURRENT APPLICATION NUMBER: US/09/930,915A
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/226,867
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,843
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 313
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-930-915A-64

Query Match 100.0%; Score 74; DB 12; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 13

US-10-082-014-145
; Sequence 145, Application US/10082014
; Publication No. US20030185858A1
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: IMMUNOGENIC Hbc CHIMER PARTICLES STABILIZED WITH AN N-
TERMINAL CYSTEINE
; FILE REFERENCE: ICC-130.0 4564/85124
; CURRENT APPLICATION NUMBER: US/10/082,014
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 09/930,915
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 145
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-082-014-145

Query Match 100.0%; Score 74; DB 12; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 14

US-10-372-076-175
; Sequence 175, Application US/10372076
; Publication No. US20030198645A1
; GENERAL INFORMATION:
; APPLICANT: Page, Mark
; APPLICANT: Friede, Martin
; TITLE OF INVENTION: STABILIZED HBc CHIMER PARTICLES AS THERAPEUTIC VACCINE
FOR
; TITLE OF INVENTION: CHRONIC HEPATITIS
; FILE REFERENCE: 4564/87179
; CURRENT APPLICATION NUMBER: US/10/372,076
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: 10/080,299
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 10/082,014
; PRIOR FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 175
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-372-076-175

Query Match 100.0%; Score 74; DB 12; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 15

US-09-865-294-3
; Sequence 3, Application US/09865294
; Publication No. US20030068325A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Immunogenic peptide composition as vaccines for the
; TITLE OF INVENTION: prevention and treatment of Alzheimer's Disease

; FILE REFERENCE: 1151-4167
; CURRENT APPLICATION NUMBER: US/09/865,294
; CURRENT FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-865-294-3

Query Match 100.0%; Score 74; DB 11; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | | |
Db 3 QYIKANSKFIGITEL 17

Search completed: December 13, 2003, 13:36:49
Job time : 2.7033 secs

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 0.915751 Seconds
(without alignments)
693.052 Million cell updates/sec

Title: US-09-785-215-4
Perfect score: 74
Sequence: 1 QYIKANSKFIGITEL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	74	100.0	15	2	US-08-319-704-10	Sequence 10, Appl
2	74	100.0	15	2	US-08-661-052-6	Sequence 6, Appli
3	74	100.0	15	2	US-08-460-502-7	Sequence 7, Appli
4	74	100.0	15	3	US-09-046-373-2	Sequence 2, Appli
5	74	100.0	15	3	US-09-188-082-6	Sequence 6, Appli
6	74	100.0	15	4	US-09-364-088-6	Sequence 6, Appli
7	74	100.0	15	4	US-09-102-716-6	Sequence 6, Appli
8	74	100.0	15	4	US-09-148-711A-7	Sequence 7, Appli
9	74	100.0	15	4	US-08-716-249-4	Sequence 4, Appli
10	74	100.0	15	5	PCT-US93-11703-69	Sequence 69, Appl
11	74	100.0	16	3	US-09-248-588-55	Sequence 55, Appl

12	74	100.0	17	1	US-08-446-692-4	Sequence 4, Appli
13	74	100.0	17	2	US-08-488-351A-4	Sequence 4, Appli
14	74	100.0	17	3	US-09-100-409A-40	Sequence 40, Appl
15	74	100.0	17	4	US-08-472-701-23	Sequence 23, Appl
16	74	100.0	17	5	PCT-US95-08596-23	Sequence 23, Appl
17	74	100.0	17	5	PCT-US95-13841-7	Sequence 7, Appli
18	74	100.0	27	1	US-08-446-692-13	Sequence 13, Appl
19	74	100.0	27	2	US-08-488-351A-13	Sequence 13, Appl
20	74	100.0	31	5	PCT-US93-11703-63	Sequence 63, Appl
21	74	100.0	37	1	US-08-446-692-57	Sequence 57, Appl
22	74	100.0	37	1	US-08-446-692-63	Sequence 63, Appl
23	74	100.0	37	2	US-08-488-351A-57	Sequence 57, Appl
24	74	100.0	37	2	US-08-488-351A-63	Sequence 63, Appl
25	74	100.0	47	1	US-08-446-692-35	Sequence 35, Appl
26	74	100.0	47	2	US-08-488-351A-35	Sequence 35, Appl
27	74	100.0	50	4	US-08-945-289-8	Sequence 8, Appli
28	74	100.0	853	4	US-08-913-880C-17	Sequence 17, Appl
29	74	100.0	858	4	US-08-913-880C-16	Sequence 16, Appl
30	74	100.0	860	4	US-08-913-880C-15	Sequence 15, Appl
31	74	100.0	862	4	US-08-913-880C-14	Sequence 14, Appl
32	74	100.0	865	4	US-08-913-880C-13	Sequence 13, Appl
33	74	100.0	866	4	US-08-913-880C-12	Sequence 12, Appl
34	74	100.0	874	4	US-08-913-880C-11	Sequence 11, Appl
35	74	100.0	875	4	US-08-913-880C-10	Sequence 10, Appl
36	74	100.0	1315	4	US-08-913-880C-1	Sequence 1, Appli
37	70	94.6	14	1	US-08-186-266-5	Sequence 5, Appli
38	70	94.6	14	1	US-08-305-871A-5	Sequence 5, Appli
39	70	94.6	14	1	US-08-465-167A-18	Sequence 18, Appl
40	70	94.6	14	2	US-08-817-933A-9	Sequence 9, Appli
41	70	94.6	14	4	US-08-464-496-15	Sequence 15, Appl
42	70	94.6	14	4	US-08-464-496-30	Sequence 30, Appl
43	70	94.6	14	4	US-08-788-822A-6	Sequence 6, Appli
44	70	94.6	14	4	US-08-197-484-95	Sequence 95, Appl
45	70	94.6	14	4	US-08-627-820-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1

US-08-319-704-10

; Sequence 10, Application US/08319704

; Patent No. 5814617

; GENERAL INFORMATION:

; APPLICANT: Hoffman, Stephen L.

; APPLICANT: Charoenvit, Yupin

; APPLICANT: Hedstrom, Richard C.

; APPLICANT: Doolan, Denise L.

; TITLE OF INVENTION: Protective 17 kDa Malaria Hepatic and

; TITLE OF INVENTION: Erythrocytic Stage Immunogen and Gene

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Naval Medical R & D Command

; STREET: Bldg 1, T-12, 8901 Wisconsin Avenue

; CITY: Bethesda

; STATE: Maryland

; COUNTRY: U.S.A

```

;      ZIP: 20889-5606
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.25
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/319,704
;      FILING DATE: 07-OCT-1994
;      CLASSIFICATION: 424
;      ATTORNEY/AGENT INFORMATION:
;      NAME: A. David Spevack
;      REGISTRATION NUMBER: 24,743
;      REFERENCE/DOCKET NUMBER: 75,206
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (301) 295-6759
;      TELEFAX: (301) 295-1022
;      INFORMATION FOR SEQ ID NO: 10:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 15 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: peptide
US-08-319-704-10

```

```

Query Match          100.0%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 QYIKANSKFIGITEL 15
        |||||
Db      1 QYIKANSKFIGITEL 15

```

RESULT 2

US-08-661-052-6

```

; Sequence 6, Application US/08661052
; Patent No. 5837243
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-Fc RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/661,052
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-661-052-6

```

```

Query Match          100.0%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 QYIKANSKFIGITEL 15
        |||||
Db      1 QYIKANSKFIGITEL 15

```

RESULT 3

```

US-08-460-502-7
; Sequence 7, Application US/08460502
; Patent No. 5843464
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Parvin T.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,502
; FILING DATE:
; CLASSIFICATION: 424

```

```

; ATTORNEY/AGENT INFORMATION:
;   NAME: Golrick, Mary E.
;   REGISTRATION NUMBER: 34,829
;   REFERENCE/DOCKET NUMBER: 22727/00120
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (216) 622-8458
;   TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 15 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-460-502-7

```

```

Query Match          100.0%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 QYIKANSKFIGITEL 15
        |||||
Db      1 QYIKANSKFIGITEL 15

```

RESULT 4

US-09-046-373-2

```

; Sequence 2, Application US/09046373
; Patent No. 6235714
; GENERAL INFORMATION:
;   APPLICANT: Sudhir Paul
;   APPLICANT: Larry J. Smith
;   APPLICANT: Gennady Gololobov
;   TITLE OF INVENTION: Methods for Identifying Inducers and
;   TITLE OF INVENTION: Inhibitors of Catalytic Antibodies, Compositions and
Their
;   TITLE OF INVENTION: Use
;   FILE REFERENCE: UNMC 63123
;   CURRENT APPLICATION NUMBER: US/09/046,373
;   CURRENT FILING DATE: 1998-03-23
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Clostridium tetani
US-09-046-373-2

```

```

Query Match          100.0%; Score 74; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 QYIKANSKFIGITEL 15
        |||||
Db      1 QYIKANSKFIGITEL 15

```

RESULT 5

US-09-188-082-6

; Sequence 6, Application US/09188082
 ; Patent No. 6270765
 ; GENERAL INFORMATION:
 ; APPLICANT: Yashwant M. Deo
 ; APPLICANT: Joel Goldstein
 ; APPLICANT: Robert Graziano
 ; APPLICANT: Chezian Somasundaram
 ; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
 ; TITLE OF INVENTION: OF ANTI-Fc RECEPTOR ANTIBODIES
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street, Suite 510
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1875
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/188,082
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/661,052
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Arnold, Beth E.
 ; REGISTRATION NUMBER: 35,430
 ; REFERENCE/DOCKET NUMBER: MXI-043CP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)227-5941
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 15 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: internal

US-09-188-082-6

Query Match 100.0%; Score 74; DB 3; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.9e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||
 Db 1 QYIKANSKFIGITEL 15

RESULT 6

US-09-364-088-6

```

; Sequence 6, Application US/09364088
; Patent No. 6365161
; GENERAL INFORMATION:
;   APPLICANT: Yashwant M. Deo, et al.
;   TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
;   TITLE OF INVENTION: OF ANTI-Fc RECEPTOR ANTIBODIES
;   NUMBER OF SEQUENCES: 16
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: LAHIVE & COCKFIELD, LLP
;     STREET: 28 State Street, 24th Floor
;     CITY: Boston
;     STATE: Massachusetts
;     COUNTRY: USA
;     ZIP: 02109
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/364,088
;     FILING DATE:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 09/188,082
;     FILING DATE: 07-JUNE-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/484,172
;     FILING DATE: 07-JUNE-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Remillard, Jane E.
;     REGISTRATION NUMBER: 38,872
;     REFERENCE/DOCKET NUMBER: MXI-043CP2
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (617)227-7400
;     TELEFAX: (617)742-7414
;   INFORMATION FOR SEQ ID NO: 6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 15 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: peptide
;     FRAGMENT TYPE: internal
US-09-364-088-6

```

```

Query Match          100.0%; Score 74; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 QYIKANSKFIGITEL 15
        |||
Db      1 QYIKANSKFIGITEL 15

```

```

RESULT 7
US-09-102-716-6
; Sequence 6, Application US/09102716
; Patent No. 6395272

```

```

; GENERAL INFORMATION:
;   APPLICANT: Yashwant M. Deo
;               Joel Goldstein
;               Robert Graziano
;               Chezian Somasundaram
;   TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
;                       OF ANTI-Fc RECEPTOR ANTIBODIES
;   NUMBER OF SEQUENCES: 16
;   CORRESPONDENCE ADDRESS:
;       ADDRESSEE: LAHIVE & COCKFIELD
;       STREET: 60 State Street, Suite 510
;       CITY: Boston
;       STATE: Massachusetts
;       COUNTRY: USA
;       ZIP: 02109-1875
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/102,716
;       FILING DATE: 22-Jun-1998
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 08/661,052
;       FILING DATE: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;       NAME: Arnold, Beth E.
;       REGISTRATION NUMBER: 35,430
;       REFERENCE/DOCKET NUMBER: MXI-043CP
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (617)227-7400
;       TELEFAX: (617)227-5941
;   INFORMATION FOR SEQ ID NO: 6:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 15 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FRAGMENT TYPE: internal
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-102-716-6

```

```

Query Match          100.0%;  Score 74;  DB 4;  Length 15;
Best Local Similarity 100.0%;  Pred. No. 1.9e-07;
Matches   15;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

```

```

Qy      1 QYIKANSKFIGITEL 15
        |||||||||||||
Db      1 QYIKANSKFIGITEL 15

```

```

RESULT 8
US-09-148-711A-7
; Sequence 7, Application US/09148711A
; Patent No. 6436405
; GENERAL INFORMATION:

```



```
; APPLICANT: The Ohio State University
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; FILE REFERENCE: 18525-04010
; CURRENT APPLICATION NUMBER: US/09/148,711A
; CURRENT FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 08/460,502
; PRIOR FILING DATE: 1995-06-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 15
; TYPE: PRT
; ORGANISM: clostridium tetani
US-09-148-711A-7
```

```
Query Match          100.0%; Score 74; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 QYIKANSKFIGITEL 15
        |||||
Db      1 QYIKANSKFIGITEL 15
```

RESULT 9

US-08-716-249-4

```
; Sequence 4, Application US/08716249
; Patent No. 6455244
; GENERAL INFORMATION:
; APPLICANT: Guichard, Gilles
; APPLICANT: Muller, Sylviane
; APPLICANT: Briand, Jean-Paul
; APPLICANT: Regenmortel, Marc
; TITLE OF INVENTION: Retropeptides, Antibodies Thereto, and
; TITLE OF INVENTION: Uses Thereof for Vaccination and In Vitro Diagnosis
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spencer & Frank
; STREET: 1100 New York Avenue, Suite 300E
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,249
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00292
; FILING DATE: 13-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Calvetti, Frederick F.
; REGISTRATION NUMBER: 28,557
```

; REFERENCE/DOCKET NUMBER: GROFO 7001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)414-4000
; TELEFAX: (202)414-4040
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-716-249-4

Query Match 100.0%; Score 74; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
|||
Db 1 QYIKANSKFIGITEL 15

RESULT 10

PCT-US93-11703-69

; Sequence 69, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0222.101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 69:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-11703-69

Query Match 100.0%; Score 74; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 11
US-09-248-588-55
; Sequence 55, Application US/09248588
; Patent No. 6231864
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and
; TITLE OF INVENTION: their Derivatives
; FILE REFERENCE: SYN-101 4564/69529
; CURRENT APPLICATION NUMBER: US/09/248,588
; CURRENT FILING DATE: 1999-02-11
; EARLIER APPLICATION NUMBER: 60/074537
; EARLIER FILING DATE: 1998-02-12
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Clostridium tetani
; PUBLICATION INFORMATION:
; JOURNAL: Vaccine
; VOLUME: 15
; ISSUE: 4
; PAGES: 377-
; DATE: 1997
US-09-248-588-55

Query Match 100.0%; Score 74; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 12
US-08-446-692-4
; Sequence 4, Application US/08446692
; Patent No. 5759551

```

; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for
vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-4

```

```

Query Match          100.0%; Score 74; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 QYIKANSKFIGITEL 15
        |||||
Db      3 QYIKANSKFIGITEL 17

```

RESULT 13

US-08-488-351A-4

```

; Sequence 4, Application US/08488351A
; Patent No. 5843446

```

```

; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

```

```

; TITLE OF INVENTION: and synthetic universal immune stimulators for
vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-4

```

```

Query Match          100.0%; Score 74; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 QYIKANSKFIGITEL 15
        |||||
Db      3 QYIKANSKFIGITEL 17

```

```

RESULT 14
US-09-100-409A-40

```

```

; Sequence 40, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-409A-40

```

```

Query Match          100.0%; Score 74; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 QYIKANSKFIGITEL 15
        |||||
Db      3 QYIKANSKFIGITEL 17

```

RESULT 15

US-08-472-701-23

```

; Sequence 23, Application US/08472701
; Patent No. 6509165
; GENERAL INFORMATION:
; APPLICANT: Griffin, Ann C.
; APPLICANT: Hickey, William F.
; TITLE OF INVENTION: Detection and Treatment Methods for

```

```

; TITLE OF INVENTION: Type I Diabetes
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,701
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,220
; FILING DATE: 08-JULY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: DCI-092DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-472-701-23

```

```

Query Match          100.0%; Score 74; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 QYIKANSKFIGITEL 15
        |||||
Db      3 QYIKANSKFIGITEL 17

```

```

Search completed: December 13, 2003, 13:31:53
Job time : 1.91575 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 2.49084 Seconds
(without alignments)
955.861 Million cell updates/sec

Title: US-09-785-215-4
Perfect score: 74
Sequence: 1 QYIKANSKFIGITEL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	74	100.0	15	11	AAR06310	Tetanus toxin epit	
2	74	100.0	15	18	AAW35506	Universal T-cell e	
3	74	100.0	15	18	AAW11505	Tetanus toxoid uni	
4	74	100.0	15	19	AAW67033	Tetanus toxin frag	
5	74	100.0	15	19	AAW71321	Universal helper T	
6	74	100.0	15	20	AAW04051	T-Helper epitope f	
7	74	100.0	15	20	AAW67578	T-cell epitope pep	
8	74	100.0	15	20	AAW73220	Tetanus toxoid epi	
9	74	100.0	15	21	AAB45511	Tetanus P2 epitope	
10	74	100.0	15	21	AAW82637	Tetanus toxoid T c	
11	74	100.0	15	21	AAW92625	Foreign epitope P2	
12	74	100.0	15	21	AAW84427	Amino acid sequenc	
13	74	100.0	15	21	AAW70300	Clostridium tetani	
14	74	100.0	15	21	AAW44763	Tetanus toxoid pro	
15	74	100.0	15	22	AAE11763	Clostridium tetani	
16	74	100.0	15	22	AAM99515	Vaccine related MH	
17	74	100.0	15	22	AAB85701	Amino acid sequenc	
18	74	100.0	15	22	AAB85451	Wild-type TT830 (t	
19	74	100.0	15	22	AAB61956	Tetanus Toxoid uni	
20	74	100.0	15	22	AAB20143	Tetanus toxin T-ce	
21	74	100.0	15	22	AAB68636	HER-2 B cell pepti	
22	74	100.0	15	22	AAB46172	Tetanus toxoid TT8	
23	74	100.0	15	22	AAB49071	Tetanus toxoid TT	
24	74	100.0	15	23	ABG31774	T helper cell epit	
25	74	100.0	15	23	AAU97872	Tetanus toxin P2 (
26	74	100.0	15	24	ABP72694	Tetanus toxoid T c	
27	74	100.0	15	24	ABG72721	Tetanus toxin immu	
28	74	100.0	16	18	AAW35445	T-cell stimulatory	
29	74	100.0	16	20	AAW29705	Clostridium tetani	
30	74	100.0	16	23	AAU93865	Clostridium tetani	
31	74	100.0	16	23	AAU11413	Tetanus toxoid pre	
32	74	100.0	17	15	AAR62692	Helper T cell epit	
33	74	100.0	17	16	AAR82573	Tetanus toxin help	
34	74	100.0	17	17	AAW05599	Tetanus toxin help	
35	74	100.0	17	17	AAR88395	T-cell antigen TT2	
36	74	100.0	17	21	AAW99274	HLA class II bindi	
37	74	100.0	17	21	AAW80056	Pathogen derived T	
38	74	100.0	17	21	AAW54539	T helper cell (Th)	
39	74	100.0	17	21	AAW58768	Unidentified pepti	
40	74	100.0	17	22	AAM99516	Vaccine related MH	
41	74	100.0	17	22	AAG62904	Amino acid residue	
42	74	100.0	17	22	AAB84435	Amino acid sequenc	
43	74	100.0	17	22	AAB30941	Amino acid sequenc	
44	74	100.0	17	22	AAB31029	Antigenic fragment	
45	74	100.0	17	22	AAB31118	Antigenic fragment	

ALIGNMENTS

RESULT 1

AAR06310

ID AAR06310 standard; protein; 15 AA.

XX

AC AAR06310;

XX

DT 04-DEC-1990 (first entry)

XX

DE Tetanus toxin epitope.

XX

KW Tetanus toxin; vaccine; major histocompatibility complex; MHC;
KW antimalarial.

XX

OS Synthetic.

XX

PN EP378881-A.

XX

PD 25-JUL-1990.

XX

PF 27-DEC-1989; 89EP-0203318.

XX

PR 16-NOV-1989; 89IT-0022409.

PR 17-JAN-1989; 89IT-0019110.

XX

PA (ENIE) ENIRICERCHE SPA.

XX

PI Pessi A, Bianchi E, Verdini AS, Corradin G;

XX

DR WPI; 1990-225582/30.

XX

PT Synthetic peptide(s) corresp. to tetanus toxin epitope(s) - used
PT as universal carriers for prepn. of immunogenic conjugate(s) for
PT use as vaccines.

XX

PS Claim 1; Page 17; 20pp; English.

XX

CC Epitopic peptides may be used with synthetic hapten derived from
CC a pathogen to generate an immune response to the pathogen.

CC Peptides are recognised by numerous T-helper cell clones within
CC the context of a wide range of alleles of the human MHC.

CC The peptides may be used in an antimalarial vaccine inducing Ab.

CC response to P.falciparum.

XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 11; Length 15;

Best Local Similarity 100.0%; Pred. No. 4.6e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15

|||||||

Db 1 QYIKANSKFIGITEL 15

RESULT 2

AAW35506

ID AAW35506 standard; peptide; 15 AA.

XX
 AC AAW35506;
 XX
 DT 25-MAR-2003 (updated)
 DT 22-APR-1998 (first entry)
 XX
 DE Universal T-cell epitope peptide SEQ ID NO:8.
 XX
 KW T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
 KW scaffold; inhibition; metastasis; wound healing; solid phase.
 XX
 OS Unidentified.
 XX
 PN WO9738011-A1.
 XX
 PD 16-OCT-1997.
 XX
 PF 03-APR-1997; 97WO-DK00146.
 XX
 PR 03-APR-1996; 96DK-0000398.
 XX
 PA (PEPR-) PEPRESEARCH AS.
 XX
 PI Heegaard PMH, Jakobsen PH;
 XX
 DR WPI; 1997-512645/47.
 XX
 PT Non-dendritic peptide carrier linked to a solid phase - useful as a
 PT diagnostic agent and as a scaffold for production of chemical
 PT derivatives
 XX
 PS Example 20; Page 124; 262pp; English.
 XX
 CC A non-dendritic peptide carrier (A) has been developed which is coupled
 CC through a linker to a solid phase, forming a complex of (A)-solid phase.
 CC Where (A) comprises 10-50 amino acids capable of forming a secondary
 CC structure in a benign buffer after liberation from the solid phase, and
 CC further the (A)-solid phase complex comprises an immunogenic substance
 CC and/or an immune mediator coupled on (A). The present sequence
 CC represents a peptide used in an example from the present invention. An
 CC (A)-solid phase complex can be used as a scaffold for the production of
 CC chemical derivatives, characterised by covalently attaching molecules at
 CC attachment points. Alternatively (A) is used as a scaffold-peptide for
 CC the incorporation into an Immunostimulating Complex (Iscom) resulting an
 CC (A)-Iscom complex which is used for the chemical coupling of antigenic
 CC substances in an aqueous solution by conjugation. (A) derivatised with
 CC one or more peptides having fibronectin-, laminin- or vitronectin-like
 CC binding activities can be used for the promotion of cell-attachment to
 CC plastic surfaces, in particular to inhibit tumour growth and metastasis,
 CC and for promotion of wound healing. Also a derivatised (A) can be used
 CC for the selection of specifically-binding aptamers or as a diagnostic
 CC agent. Such diagnostic-(A) molecules could be used to detect molecules
 CC derived from or indicative of pregnancy or of a disease, such as an
 CC infectious, autoimmune or cancerous disease.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
|||
Db 1 QYIKANSKFIGITEL 15

RESULT 3

AAW11505

ID AAW11505 standard; Protein; 15 AA.

XX

AC AAW11505;

XX

DT 24-SEP-1997 (first entry)

XX

DE Tetanus toxoid universal Th epitope TT830.

XX

KW Humanised antibody; anti-Fc receptor; H22; bifunctional; bispecific;
KW fusion protein; chimera; tetanus toxoid; helper T cell epitope;
KW antigen presentation; ds.

XX

OS Clostridium tetani.

XX

PN WO9640789-A1.

XX

PD 19-DEC-1996.

XX

PF 07-JUN-1996; 96WO-US09988.

XX

PR 07-JUN-1995; 95US-0484172.

XX

PA (MEDA-) MEDAREX INC.

XX

PI Deo YM, Goldstein J, Graziano R, Somasundaram C;

XX

DR WPI; 1997-052242/05.

DR N-PSDB; AAT58127.

XX

PT Recombinant, multi-specific anti-Fc receptor antibody molecules -
PT also comprise an anti-target portion, used for the treatment of
PT cancer, autoimmune disease and pathogenic infection

XX

PS Example 7; Fig 24; 115pp; English.

XX

CC Synthetic DNA coding for the wild-type universal Th epitope from
CC tetanus toxoid, designated TT830, was fused to the 3'-end of DNA
CC encoding heavy chain sequences from the humanised anti-Fc gamma RI
CC monoclonal antibody H22. The resulting fusion protein was shown to
CC be significantly more efficient in antigen presentation and T cell
CC stimulation than the TT830 epitope alone. A similar fusion
CC construct was prepared coding for a mutant, antagonistic form of the
CC epitope (designated TT833S) fused to the anti-Fc gamma RI. The
CC Fab22-TT833S is at least 100 times more effective than TT833S in
CC inhibiting T cell activation.

XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 18; Length 15;

Best Local Similarity 100.0%; Pred. No. 4.6e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15

|||||||

Db 1 QYIKANSKFIGITEL 15

RESULT 4

AAW67033

ID AAW67033 standard; peptide; 15 AA.

XX

AC AAW67033;

XX

DT 15-DEC-1998 (first entry)

XX

DE Tetanus toxin fragment (residues 830-844).

XX

KW Tetanus toxin; vaccine; antibody; carbohydrate peptide conjugate;
KW dendrimeric poly-lysine; epitope; tumour.

XX

OS Clostridium tetani.

XX

PN WO9843677-A1.

XX

PD 08-OCT-1998.

XX

PF 27-MAR-1998; 98WO-EP01922.

XX

PR 27-MAR-1997; 97US-0041726.

XX

PA (INSP) INST PASTEUR.

XX

PI Bay S, Cantacuzene D, Leclerc C, Lo-man R;

XX

DR WPI; 1998-557071/47.

XX

PT Carbohydrate peptide conjugate used as vaccine - comprises carrier
PT with dendrimeric poly-lysine enabling multiple epitopes to be
PT covalently attached

XX

PS Disclosure; Page 13; 55pp; English.

XX

CC The invention relates to a new carbohydrate peptide conjugate, which
CC comprises a carrier with a dendrimeric poly-lysine enabling multiple
CC epitopes to be covalently attached to it. Also claimed are: (1) an
CC antibody purified from biological fluid or cells of organisms
CC administered with the carbohydrate peptide conjugate, and (2) a diagnosis
CC kit comprising antigen-specific antibodies elicited by immunisation with
CC the carbohydrate peptide conjugate. The peptide conjugate, antibody and
CC diagnosis kit are used to provide pharmaceutical compositions and
CC vaccines against tumours. These can be used to support an immune response
CC against viral infections caused by hepatitis virus, HIV or cytomegalo

CC virus. They can be used to enhance immune responses, especially B- and T-
CC cell responses, of humans and animals against bacterial infections. The
CC carbohydrate peptide conjugate stimulates the antibody and T-cell
CC response without stimulating undesired immune responses. The composition
CC is capable of increasing the survival of tumour bearing humans and
CC animals. The present sequence corresponds to residues 830-844 of tetanus
CC toxin. The synthetic peptide corresponding to this sequence may be used
CC as an epitope in a carbohydrate peptide conjugate.

XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 19; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15

|||||||

Db 1 QYIKANSKFIGITEL 15

RESULT 5

AAW71321

ID AAW71321 standard; peptide; 15 AA.

XX

AC AAW71321;

XX

DT 26-NOV-1998 (first entry)

XX

DE Universal helper T-cell epitope P2 derived from tetanus toxin.

XX

KW Liver stage; Plasmodium; Navy Yoelii Liver Stage 3 antibody; NYLS3;

KW hepatic and erythrocytic stage protein; PyHEP17; vaccine;

KW malaria parasite; tetanus toxin; P2; helper T-cell epitope.

XX

OS Synthetic.

OS Clostridium tetani.

XX

PN US5814617-A.

XX

PD 29-SEP-1998.

XX

PF 07-OCT-1994; 94US-0319704.

XX

PR 07-OCT-1994; 94US-0319704.

XX

PA (USNA) US SEC OF NAVY.

XX

PI Charoenvit Y, Doolan DL, Hedstrom RC, Hoffman SL;

XX

DR WPI; 1998-541794/46.

XX

PT Vaccine for protecting mammal against infection by malaria caused by

PT Plasmodium species - comprises a first nucleic acid encoding a first

PT polypeptide capable of eliciting an immune reaction against an

PT antigen expressed during the liver

XX

PS Disclosure; Column 12; 24pp; English.

XX
 CC AAW71321-22 represent universal helper T-cell epitopes derived from
 CC tetanus toxin. They are used to enhance host immune response to
 CC vaccines. The specification describes a Plasmodium yoelii liver stage
 CC 17 kDa hepatic and erythrocytic stage protein designated PyHEP17. This
 CC protein elicits a response from an Ig1 monoclonal antibody designated
 CC Navy Yoelii Liver Stage 3 (NYLS3). This antibody does not recognise
 CC sporozoites, but does recognise P. yoelii liver stage parasites. NYLS3
 CC eliminates upto 90% of liver stage parasites. The specification
 CC describes a vaccine for reducing the severity or incidence of infection
 CC by a malaria parasite of the genus Plasmodium. The DNA vaccine comprises
 CC exon 1 and part of exon 2 of the PyHEP17 gene.

XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 19; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||
 Db 1 QYIKANSKFIGITEL 15

RESULT 6

AAAY04051

ID AAY04051 standard; peptide; 15 AA.

XX

AC AAY04051;

XX

DT 04-JAN-2000 (first entry)

XX

DE T-Helper epitope from tetanus toxoid.

XX

KW Covalently reactive antigen analog; CRAA; catalytic antibody;
 KW electrophilic reaction centre; phosphonate; boronate; vaccine;
 KW transition state analog; TSA; isostere; gp120; HIV-1; T-helper;
 KW tetanus; toxoid; B-T-epitope.

XX

OS Clostridium tetani.

XX

PN WO9948925-A1.

XX

PD 30-SEP-1999.

XX

PF 23-MAR-1999; 99WO-US06325.

XX

PR 23-MAR-1998; 98US-0046373.

XX

PA (UYNE-) UNIV NEBRASKA.

XX

PI Paul S, Gololobov G, Smith L;

XX

DR WPI; 1999-591076/50.

XX

PT New covalently reactive antigen analogs used for treating e.g.
 PT autoimmune diseases, lymphoproliferative disorders, cancers, microbial

PT infections, ischemic and reperfusion injury or septic shock -
 XX
 PS Disclosure; Page 86; 158pp; English.
 XX
 CC The patent discloses new covalently reactive antigen analogs (CRAA)
 CC of formula X1-Y-E-X2, in which X1 and X2 represent peptide sequences
 CC of an epitope of a disease-associated protein, Y is a positively
 CC charged amino acid residue, preferably Lys or Arg, and E is an
 CC electrophilic reaction centre, preferably a phosphonate or boronate
 CC moiety. Depending on the identity of the epitope, the CRAA may be used
 CC to stimulate production of catalytic antibodies specific for
 CC predetermined antigens associated with particular medical disorders.
 CC They may also be used to permanently inactivate endogenously produced
 CC catalytic antibodies produced in certain autoimmune diseases as well as
 CC in certain lymphoproliferative disorders.
 CC Amongst the specifically exemplified CRAAs is one based on residues
 CC 421-436 of a B-cell epitope of gp120 (see AAY04046) which may be used
 CC to counter HIV-1 infections. When used as an immunogen, preferably this
 CC CRAA is conjugated at its N-terminal to a T-helper epitope from tetanus
 CC toxoid. The present sequence represents the T-helper epitope and
 CC corresponds to residues 830-844 of the toxoid.
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 20; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||
 Db 1 QYIKANSKFIGITEL 15

RESULT 7
 AAW67578
 ID AAW67578 standard; peptide; 15 AA.
 XX
 AC AAW67578;
 XX
 DT 02-MAR-1999 (first entry)
 XX
 DE T-cell epitope peptide #4 for chimeric fimbrin/T-cell epitope peptide.
 XX
 KW Chimeric; non-typable Haemophilus influenzae; fimbrin; T-cell epitope;
 KW immunogenic composition; immune response.
 XX
 OS Synthetic.
 XX
 PN US5843464-A.
 XX
 PD 01-DEC-1998.
 XX
 PF 02-JUN-1995; 95US-0460502.
 XX
 PR 02-JUN-1995; 95US-0460502.
 XX
 PA (OHIS) UNIV OHIO STATE.

XX
 PI Bakaletz LO, Kaumaya PTP;
 XX
 DR WPI; 1999-044514/04.
 XX
 PT Synthetic chimeric fimbrin peptide - useful for vaccination against
 PT non-typable Haemophilus influenzae
 XX
 PS Disclosure; Column 4; 16pp; English.
 XX
 CC The invention relates to the manufacture of a synthetic chimeric peptide
 CC comprising a non-typable Haemophilus influenzae fimbrin peptide fused via
 CC a linker peptide to a T-cell epitope peptide. The chimeric peptide is
 CC used in immunogenic compositions which induce an immune response against
 CC non-typable Haemophilus influenzae. This sequence represents an example
 CC of a T-cell epitope peptide used to generate the chimeric peptide.
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 20; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||
 Db 1 QYIKANSKFIGITEL 15

RESULT 8

AAW73220

ID AAW73220 standard; Protein; 15 AA.

XX

AC AAW73220;

XX

DT 25-JAN-1999 (first entry)

XX

DE Tetanus toxoid epitope.

XX

KW Multispecific single chain antibody; antibody H22; tumour cell; therapy;
 KW antibody-dependent cellular cytotoxicity; ADCC; HER 2/neu; infection;
 KW epidermal growth factor receptor; breast cancer; ovarian cancer.

XX

OS Synthetic.

XX

PN US5837243-A.

XX

PD 17-NOV-1998.

XX

PF 07-JUN-1996; 96US-0661052.

XX

PR 07-JUN-1996; 96US-0661052.

PR 07-JUN-1995; 95US-0484172.

XX

PA (MEDA-) MEDAREX INC.

XX

PI Deo YM, Goldstein J, Graziano R, Somasundaram C;

XX

DR WPI; 1999-023374/02.
 XX
 PT Specific killing of tumour cells - using a multi-specific molecule
 PT comprising an anti-Fc receptor antibody and a portion which binds to
 PT a target cell
 XX
 PS Example 7; Column 27; 57pp; English.
 XX
 CC This sequence represents a tetanus toxoid epitope and is recognised
 CC by the multispecific single chain antibody designated H22. The
 CC antibody can be used in the method of the invention for inducing
 CC antibody-dependent cellular cytotoxicity (ADCC) against a tumour cell
 CC which is characterised by overexpression of HER 2/neu or epidermal growth
 CC factor receptor (EGFR), comprises contacting the tumour cell with a
 CC multispecific protein molecule (preferably a single chain antibody)
 CC comprising: (a) an anti-Fc receptor antibody or an antigen binding
 CC fragment; (b) a portion which binds to HER 2/neu; and (c) a portion which
 CC binds to EGFR. The method can be used for treating cancers especially
 CC breast cancer or ovarian cancer. The multispecific antibody can also
 CC be administered prophylactically to vaccinate a subject against infection
 CC by a target cell.
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 20; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||
 Db 1 QYIKANSKFIGITEL 15

RESULT 9
 AAB45511
 ID AAB45511 standard; Protein; 15 AA.
 XX
 AC AAB45511;
 XX
 DT 26-FEB-2001 (first entry)
 XX
 DE Tetanus P2 epitope SEQ ID NO: 23.
 XX
 KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
 KW cancer; eosinophilia; vaccine; allergic rhinitis.
 XX
 OS Clostridium tetani.
 XX
 PN WO200065058-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 19-APR-2000; 2000WO-DK00205.
 XX
 PR 23-APR-1999; 99DK-0000552.
 PR 06-MAY-1999; 99US-0132811.
 XX

PA (MEBI-) M & E BIOTECH AS.

XX

PI Klysner S;

XX

DR WPI; 2000-672791/65.

XX

PT Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -

XX

PS Example 1; Page 137; 172pp; English.

XX

CC The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections.

XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 21; Length 15;

Best Local Similarity 100.0%; Pred. No. 4.6e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15

|||||||

Db 1 QYIKANSKFIGITEL 15

RESULT 10

AAAY82637

ID AAY82637 standard; peptide; 15 AA.

XX

AC AAY82637;

XX

DT 07-AUG-2000 (first entry)

XX

DE Tetanus toxoid T cell epitope peptide SEQ ID NO:13.

XX

KW T cell epitope; B cell epitope; allergy; allergen; antigenic;
KW antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KW atopic dermatitis; acute urticaria; chronic urticaria;
KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
KW anaphylactic reaction; drug hypersensitivity; allergic reaction.

XX

OS Clostridium tetani.

OS Synthetic.

XX

PN WO200006694-A2.

XX

PD 10-FEB-2000.

XX

PF 20-JUL-1999; 99WO-BE00092.

XX

PR 30-JUL-1998; 98EP-0870167.

XX

PA (UNIO) UCB SA.

XX

PI Saint-Remy J, Jacquemin M;

XX

DR WPI; 2000-422470/36.

XX

PT New compound for prevention and treatment of allergies comprises at
PT least one allergen antigenic determinant recognized by a B cell and at
PT least one antigenic determinant which does not trigger T cell
PT activation -

XX

PS Example 6; Page 30; 50pp; English.

XX

CC The present invention describes a compound (I) for the prevention and/or
CC treatment of allergy. The compound comprises at least one allergen
CC antigenic determinant (i) recognised by a B cell or an antibody secreted
CC by a B cell of a non-atopic individual and at least one antigenic
CC determinant (ii) different from the allergen that triggers T cell
CC activation. (I) has antiallergic, antiasthmatic, antiinflammatory.
CC dermatological and immunosuppressive activities, and can be used in a
CC vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to
CC treat and/or prevent allergies or a disease of allergic origin,
CC especially hypersensitivities. These include rhinitis, sinusitis,
CC bronchial asthma, atopic dermatitis, some forms of acute and chronic
CC urticaria, gastro-intestinal syndromes associated with the ingestion of
CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions
CC associated with drug hypersensitivities and/or a mixture of these. The
CC use of (I) in the treatment of allergic conditions avoids the need for
CC drug treatment, which often causes undesirable side-effects. Also, prior
CC art drug therapies alleviate symptoms, but do not influence their
CC causes, however (I) actually combats the cause of an allergic reaction.
CC The present sequence represents a peptide, which is used in an
CC example from the present invention.

XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 21; Length 15;

Best Local Similarity 100.0%; Pred. No. 4.6e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15

|||||||

Db 1 QYIKANSKFIGITEL 15

RESULT 11

AA92625

ID AA92625 standard; Protein; 15 AA.

XX

AC AA92625;

XX

DT 10-AUG-2000 (first entry)

XX

DE Foreign epitope P2.
XX
KW Foreign epitope; P2; prostate specific membrane antigen; PSM; Her2;
KW Heregulin 2; Fibroblast growth factor 8b; FGF8b; vaccination;
KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
KW prostate cancer; cell-associated peptide antigen.
XX
OS Clostridium tetani.
XX
PN WO200020027-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-DK00525.
XX
PR 05-OCT-1998; 98DK-0001261.
PR 20-OCT-1998; 98US-0105011.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
DR WPI; 2000-349917/30.
DR N-PSDB; AAA09460.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
PS Example 1; Page 213; 220pp; English.
XX
CC The claims detail a method for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, human
CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope (e.g. P2 and/or P30) are also claimed.
CC The method is used to treat prostate, prostate/breast or breast cancer
CC when the PA is human PSM, FGF8b and Her2, respectively.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
|||||||
Db 1 QYIKANSKFIGITEL 15

RESULT 12

AAAY84427

ID AAY84427 standard; Peptide; 15 AA.

XX

AC AAY84427;

XX

DT 25-JUL-2000 (first entry)

XX

DE Amino acid sequence of the tetanus toxoid P2 epitope.

XX

KW Osteoprotegerin ligand; OPGL; osteoprotegerin; osteoclastogenesis;

KW tumour necrosis factor receptor; type II transmembrane protein;

KW osteoclast differentiation; CSF-1; osteoclast activator;

KW immune response; osteoporosis; bone resorption;

KW tetanus toxoid P2 epitope.

XX

OS Clostridium tetani.

XX

PN WO200015807-A1.

XX

PD 23-MAR-2000.

XX

PF 13-SEP-1999; 99WO-DK00481.

XX

PR 15-SEP-1998; 98DK-0001164.

PR 02-OCT-1998; 98US-0102896.

XX

PA (MEBI-) M & E BIOTECH AS.

XX

PI Halkier T, Haaning J;

XX

DR WPI; 2000-271444/23.

XX

PT In vivo down-regulation of osteoprotegerin ligand (OPGL) activity used
PT to treat, prevent and ameliorate osteoporosis -

XX

PS Example; Page 106; 110pp; English.

XX

CC The present sequence represents the tetanus toxoid P2 epitope. It is

CC used to create a fusion protein with murine osteoprotegerin ligand

CC (OPGL). Osteoprotegerin is a secreted member of the tumour necrosis

CC factor receptor family, which blocks osteoclastogenesis in a dose

CC dependent manner. The OPGL protein is synthesised as a type II

CC transmembrane protein. The murine and human OPGL polypeptides are 87%

CC homologous. OPGL is a potent osteoclast differentiation factor when

CC combined with CSF-1. It is not capable of inducing osteoclast

CC differentiation in the absence of CSF-1. OPGL is also an activator of

CC mature osteoclasts. The specification describes a method for the in vivo

CC down-regulation of OPGL activity in an animal. The method comprises

CC using at least one OPGL polypeptide or subsequence, and/or at least one

CC OPGL analogue to induce an immune response in the animal. The method

CC and OPGL polypeptide are useful for treating, preventing and ameliorating

CC osteoporosis or other diseases or conditions characterised by excessive

CC bone resorption.

XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||||
Db 1 QYIKANSKFIGITEL 15

RESULT 13

AAAY70300

ID AAY70300 standard; peptide; 15 AA.

XX

AC AAY70300;

XX

DT 06-JUN-2000 (first entry)

XX

DE Clostridium tetani tetanus toxoid T-cell epitope, P589.

XX

KW Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;

KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;

KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;

KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;

KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;

KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;

KW Pfg27; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody.

XX

OS Clostridium tetani.

XX

PN WO200011179-A1.

XX

PD 02-MAR-2000.

XX

PF 19-AUG-1999; 99WO-US18869.

XX

PR 21-AUG-1998; 98US-0097703.

XX

PA (NAIM-) NAT INST IMMUNOLOGY.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Lal AA, Shi YP, Hasnain SE;

XX

DR WPI; 2000-237654/20.

XX

PT Novel recombinant protein as vaccine for treating malarial infection

PT comprises antigenic peptides obtained from different stages of

PT plasmodium falciparum life cycle -

XX

PS Claim 2; Page 17; 52pp; English.

XX

CC The present sequence is the tetanus toxoid T-cell epitope P589, derived

CC from Clostridium tetani. It is used in the construction of recombinant

CC protein CDC/NIIMALVAC-1, which is a multivalent, multistage malarial

CC vaccine. The recombinant protein comprises, melittin signal peptide,

CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes

CC from circumsporozoite protein (CSP), sporozoite surface protein-2

CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
 CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
 CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
 CC specific antigen, Pfg27. These epitopes were obtained at different stages
 CC of the life cycle of P. falciparum. CDC/NIIMALVAC-1 vaccine has
 CC antiparasitic activity and can be used for treatment and prevention of
 CC malarial infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for
 CC detecting P. falciparum in biological samples.

XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||
 Db 1 QYIKANSKFIGITEL 15

RESULT 14

AAAY44763

ID AAY44763 standard; peptide; 15 AA.

XX

AC AAY44763;

XX

DT 04-MAY-2000 (first entry)

XX

DE Tetanus toxoid protein derived T-cell activating epitope P2.

XX

KW Surface layer protein; S-layer secretion signal; antibiotic; vaccine;
 KW recombinant fusion protein cleavage; enzyme; protein polymer; foodstuff;
 KW antibacterial enzyme; surface glycoprotein; T-cell activating epitope;
 KW P2; tetanus toxoid; IPNV; Infectious pancreatic necrosis virus.

XX

OS Clostridium tetani.

XX

PN WO200004170-A1.

XX

PD 27-JAN-2000.

XX

PF 14-JUL-1999; 99WO-CA00637.

XX

PR 14-JUL-1998; 98CA-2237704.

XX

PA (UYBR-) UNIV BRITISH COLUMBIA.

XX

PI Smit J;

XX

DR WPI; 2000-182434/16.

XX

PT Cleavage of Caulobacter produced recombinant fusion proteins useful for
 PT producing vaccine peptides -

XX

PS Example 2; Page 16; 33pp; English.

XX

CC The patent discloses a method for cleaving a recombinant fusion protein

CC which is produced by Caulobacter and consists of Caulobacter surface
 CC layer (S-layer) protein (containing the C-terminal secretion signal) and
 CC a target protein heterologous to Caulobacter. The cleavage of target
 CC protein from the S-layer protein is carried out under mild acid
 CC conditions so that cleavage occurs at aspartate-proline dipeptide site
 CC without solubilising the protein. The cleavage is accomplished while the
 CC fusion protein is in an insoluble aggregate form which facilitates
 CC purification of the protein. The method is useful for producing pure
 CC proteins including recombinant human and animal therapeutic antibiotic
 CC and vaccine peptides, enzymes, protein polymers, and antibacterial
 CC enzymes for foodstuffs.
 CC The present sequence is a T-cell activating epitope P2 derived
 CC from tetanus toxoid protein. This sequence was fused to a DNA encoding
 CC a fragment of Infectious pancreatic necrosis virus surface glycoprotein
 CC which is a vaccine candidate. This chimeric protein was in turn fused to
 CC DNA encoding C. crescentus S-layer secretion signal (corresponds to the
 CC C-terminal portion of the S-layer protein from amino acid 690 onwards and
 CC contains native Asp-Pro site) for construction of a recombinant
 CC fusion construct which is expressed in Caulobacter and then cleaved
 CC to recover the vaccine candidate protein.

XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
 |||||
 Db 1 QYIKANSKFIGITEL 15

RESULT 15

AAE11763

ID AAE11763 standard; peptide; 15 AA.

XX

AC AAE11763;

XX

DT 18-DEC-2001 (first entry)

XX

DE Clostridium tetani P2 epitope.

XX

KW Amyloid protein; neuroprotective; nootropic; immunostimulant; vaccine;

KW Alzheimer's disease; anticonvulsant; gene therapy; Pick's disease;

KW antidiabetic; systemic amyloidosis; maturity onset diabetes; ALS;

KW amyotrophic lateral sclerosis; Parkinson's disease; encephalopathy;

KW Huntington's disease; fronto-temporal dementia; P2 epitope.

XX

OS Clostridium tetani.

XX

PN WO200162284-A2.

XX

PD 30-AUG-2001.

XX

PF 19-FEB-2001; 2001WO-DK00113.

XX

PR 21-FEB-2000; 2000DK-0000265.

PR 01-MAR-2000; 2000US-186295P.

XX

PA (MEBI-) M & E BIOTECH AS.

XX

PI Birk P, Jensen MR, Nielsen KG;

XX

DR WPI; 2001-589796/66.

DR N-PSDB; AAD18755.

XX

PT In vivo down-regulation of amyloid protein for the treatment of
PT Alzheimer's, comprises presenting an amyloidogenic polypeptide or its
PT subsequence and/or at least one analogue of the amyloidogenic
PT polypeptide to the immune system -

XX

PS Example 3; Page 117; 120pp; English.

XX

CC The invention relates to a method for in vivo down-regulation of amyloid
CC protein such as beta amyloid (Abeta) in an animal, including human. The
CC method comprising presenting to the animal's immune system an
CC immunogenically effective amount of at least one amyloidogenic protein
CC or its subsequence and/or at least one analogue of the amyloidogenic
CC polypeptide. The amyloidogenic protein or its subsequence, and its
CC analogue is useful for the preparation of an immunogenic composition
CC comprising an adjuvant for down-regulating amyloid in an animal. They are
CC also useful in the treatment, prophylaxis or amelioration of Alzheimer's
CC disease or other diseases characterised by amyloid deposits. They are
CC also useful in the treatment of systemic amyloidosis, maturity onset
CC diabetes, Parkinson's disease, Huntington's disease, fronto-temporal
CC dementia, amyotrophic lateral sclerosis (ALS), Pick's disease and
CC prion-related transmissible spongiform encephalopathies. They are also
CC useful for inducing production of antibodies against an amyloidogenic
CC polypeptide. The present sequence is Clostridium tetani P2 epitope
CC related to the invention.

XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 74; DB 22; Length 15;

Best Local Similarity 100.0%; Pred. No. 4.6e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15

||||||||||||||

Db 1 QYIKANSKFIGITEL 15

Search completed: December 13, 2003, 13:34:23

Job time : 3.49084 secs

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 3.48718 Seconds
(without alignments)
955.861 Million cell updates/sec

Title: US-09-785-215-6
Perfect score: 112
Sequence: 1 FNNFTVSFWLRVPKVSASHLE 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	112	100.0	21	12	AAR11896		Immunogenic conjug
2	112	100.0	21	17	AAW06130		Tetanus toxoid pro
3	112	100.0	21	17	AAR88397		T-cell antigen TT3
4	112	100.0	21	18	AAW46449		Broad range helper
5	112	100.0	21	19	AAW67034		Tetanus toxin frag
6	112	100.0	21	20	AAW67579		T-cell epitope pep
7	112	100.0	21	20	AAW73222		Tetanus toxoid epi
8	112	100.0	21	21	AAB45512		Tetanus P30 epitop
9	112	100.0	21	21	AAV99876		Tetanus toxin T ce
10	112	100.0	21	21	AAV92626		Foreign epitope P3
11	112	100.0	21	21	AAV84428		Amino acid sequenc
12	112	100.0	21	21	AAV49260		CD4+ T cell epitop
13	112	100.0	21	22	AAE11764		Clostridium tetani
14	112	100.0	21	22	AAB85702		Amino acid sequenc
15	112	100.0	21	22	AAB85453		Universal tetanus
16	112	100.0	21	22	AAB61958		Tetanus Toxoid uni
17	112	100.0	21	22	AAB20144		Tetanus toxin T-ce
18	112	100.0	21	22	AAB68637		HER-2 B cell pepti
19	112	100.0	21	22	AAB46173		Tetanus toxoid TT9
20	112	100.0	21	22	AAB49072		Tetanus toxoid TT
21	112	100.0	21	23	ABG31775		T helper cell epit
22	112	100.0	21	23	AAU11415		Tetanus toxoid pre
23	112	100.0	21	24	ABP72695		Tetanus toxoid T c
24	112	100.0	28	22	AAB46176		Tetanus toxoid 947
25	112	100.0	31	21	AAV92653		PSMpep010 - P30 in
26	112	100.0	31	21	AAV92654		PSMpep011 - P30 in
27	112	100.0	31	21	AAV92655		PSMpep012 - P30 in
28	112	100.0	32	15	AAR62702		LHRH-containing im
29	112	100.0	33	22	AAB49075		Amyloid beta/tetan
30	112	100.0	34	23	AAU11421		Synthetic immunoge
31	112	100.0	36	22	AAG63662		Peptide comprising
32	112	100.0	36	22	AAG63515		A peptide which ma
33	112	100.0	37	23	AAU11425		Synthetic immunoge
34	112	100.0	43	22	AAB46177		Tetanus toxoid 830
35	112	100.0	43	22	AAB49076		Amyloid beta/tetan
36	112	100.0	44	22	AAB46194		Tetanus toxoid epi
37	112	100.0	44	22	AAB49090		Amyloid beta/tetan
38	112	100.0	50	23	AAU11429		Synthetic immunoge
39	112	100.0	51	22	AAB46195		Tetanus toxoid epi
40	112	100.0	51	22	AAB49091		Amyloid beta/tetan
41	112	100.0	59	22	AAG63661		Peptide comprising
42	112	100.0	59	22	AAG63513		A peptide which ma
43	112	100.0	63	12	AAR14263		Immunogenic branch
44	112	100.0	64	12	AAR14261		Immunogenic branch
45	112	100.0	65	12	AAR14265		Immunogenic branch

ALIGNMENTS

RESULT 1

AAR11896

ID AAR11896 standard; peptide; 21 AA.

XX

AC AAR11896;

XX

DT 25-MAR-2003 (updated)

DT 19-JUL-1991 (first entry)

XX

DE Immunogenic conjugate constituent peptide, TT3.

XX

KW Malaria vaccine; major histocompatibility complex.

XX

OS Tetanus toxin.

XX

FH Key Location/Qualifiers

FT Peptide 1..14

FT /label= active fragment (claimed)

XX

PN EP427347-A.

XX

PD 15-MAY-1991.

XX

PF 07-NOV-1990; 90EP-0202948.

XX

PR 10-NOV-1989; 89IT-0022355.

XX

PA (ENIE) ENIRICERCHE SPA.

XX

PI Blanchi E, Pessi A, Corradin G;

XX

DR WPI; 1991-141874/20.

XX

PT Synthetic peptide(s) used as universal carriers - for preparing

PT immunogenic conjugates used as vaccines against Plasmodium

PT falciparum

XX

PS Claim 1; page 13; 16pp; English.

XX

CC This peptide corresps. to residues 947-967 of Tetanus toxin. It can

CC be used as a universal carrier for the prepn. of an immunogenic

CC conjugate. It is covalently bound to a peptide or polysaccharide

CC hapten derived from a pathogen. This conjugate can be used as

CC a vaccine for malaria. This peptide is recognised by different T-

CC helper cell clones in association with alleles of the human MHC.

CC It contains 2 epitopes: (a) 953-967, recognised by DR5-restricted

CC clones; and (b) 947-960, recognised by all other DR and DP-

CC restricted clones.

CC (Updated on 25-MAR-2003 to correct PI field.)

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 12; Length 21;

Best Local Similarity 100.0%; Pred. No. 4.5e-12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21

Db |||||
1 FNNFTVSFWLRLVPKVSASHLE 21

RESULT 2

AAW06130

ID AAW06130 standard; Peptide; 21 AA.

XX

AC AAW06130;

XX

DT 07-FEB-1997 (first entry)

XX

DE Tetanus toxoid protein T-cell epitope.

XX

KW Cholesteryl ester transfer protein; CETP; antigen; vaccine;

KW cardiovascular disease; atherosclerosis; tetanus toxoid;

KW T-cell epitope.

XX

OS Clostridium tetani.

XX

PN WO9634888-A1.

XX

PD 07-NOV-1996.

XX

PF 01-MAY-1996; 96WO-US06147.

XX

PR 01-MAY-1995; 95US-0432483.

XX

PA (TCEL-) T CELL SCI INC.

XX

PI Rittershaus CW, Thomas LJ;

XX

DR WPI; 1996-506103/50.

XX

PT Cholesteryl ester transfer protein B cell epitope linked to T cell

PT epitope - used to generate vaccine to regulate CETP activity for

PT decreasing the risk of developing a cardiovascular disease e.g.

PT atherosclerosis

XX

PS Claim 11; Page 43; 72pp; English.

XX

CC A helper T-cell epitope (AAW06130) comprises amino acids 947-967

CC of tetanus toxoid protein. It can be utilised in novel peptide

CC vaccines (see also AAW06129, AAW06132) also including B-cell

CC epitope(s) from human or rabbit cholesteryl ester transfer

CC protein (CETP) to elicit an immune response against endogenous

CC CETP activity, thereby treating or preventing a cardiovascular

CC disease, such as atherosclerosis.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 17; Length 21;

Best Local Similarity 100.0%; Pred. No. 4.5e-12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRLVPKVSASHLE 21
 |||||

Db

1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 3

AAR88397

ID AAR88397 standard; Peptide; 21 AA.

XX

AC AAR88397;

XX

DT 12-JUN-1996 (first entry)

XX

DE T-cell antigen TT3 peptide.

XX

KW T-antigen; vaccine; antibody; T-cell; T-lymphocyte;

KW alpha-helix; coiled-coil heterodimer; core peptide; subunit.

XX

OS Synthetic.

XX

PN WO9531480-A1.

XX

PD 23-NOV-1995.

XX

PF 18-MAY-1995; 95WO-CA00293.

XX

PR 18-MAY-1994; 94US-0245507.

XX

PA (SPIS-) SPI SYNTHETIC PEPTIDES INC.

XX

PI Cachia PJ, Hodges RS, Houston ME, Irvin RT, Kay CM;

PI Zhou NE;

XX

DR WPI; 1996-010880/01.

XX

PT Hetero:dimeric polypeptide immunogen in coiled-coil configuration

PT with different antigens on each sub:unit - useful in vaccines and

PT for antibody prodn.

XX

PS Claim 7; Page 62; 95pp; English.

XX

CC This T-cell antigen TT3 peptide may be attached to a core peptide

CC contained in one of the 2 subunits of an alpha-helical coiled-coil

CC heterodimer. Each core peptide is comprised of terminal and

CC internal AA repeat sequences. This peptide antigen is attached

CC to the core peptide through covalent linkages to certain AA of the

CC internal repeats. The 2 subunits of the heterodimer are arranged

CC in a stable alpha-helical coiled-coil configuration having a 1:1

CC stoichiometry, and the peptide antigen is disposed toward the outer

CC surfaces of the configuration. The heterodimer may be used as a

CC synthetic vaccine (optionally multivalent) or to generate

CC antibodies.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 17; Length 21;

Best Local Similarity 100.0%; Pred. No. 4.5e-12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
 |||||
 Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 4

AAW46449

ID AAW46449 standard; Peptide; 21 AA.

XX

AC AAW46449;

XX

DT 18-MAY-1998 (first entry)

XX

DE Broad range helper T cell epitope from the tetanus toxoid protein.

XX

KW Cholesteryl ester transfer protein; CETP; cholesteryl ester;
 KW high density lipoprotein; HDL; very low density lipoprotein; VLDL;
 KW low density lipoprotein; LDL; T cell epitope; antibody;
 KW DNA plasmid-based vaccine; broad range helper T cell epitope;
 KW treatment; cardiovascular disease.

XX

OS Clostridium tetani.

XX

PN WO9741227-A1.

XX

PD 06-NOV-1997.

XX

PF 01-MAY-1997; 97WO-US07294.

XX

PR 21-FEB-1997; 97US-0802967.

PR 01-MAY-1996; 96US-0640713.

XX

PA (TCEL-) T CELL SCI INC.

XX

PI Thomas LJ;

XX

DR WPI; 1997-549731/50.

XX

PT DNA plasmid-based vaccine encodes CETP B cell and helper T cell
 PT epitope(s) - used for elevating high density lipoprotein levels, and
 PT for treating cardiovascular disease

XX

PS Disclosure; Page 44; 67pp; English.

XX

CC The present sequence represents a broad range helper T cell epitope
 CC of the tetanus protein. It can be used in DNA plasmid-based vaccines
 CC against cholesteryl ester transfer proteins (CETPs). CETPs mediate the
 CC transfer of cholesteryl esters from high density lipoprotein (HDL)
 CC to very low density lipoprotein (VLDL) and low density lipoprotein (LDL),
 CC and vice versa. An increased CETP activity produces an atherogenic
 CC lipoprotein profile and induces atherosclerosis. A DNA plasmid-based
 CC vaccine comprises sequences encoding at least one B cell epitope of CETP
 CC linked in frame with at least one segment encoding a broad range helper
 CC T cell epitope. The vaccines can be used to elevate the ratio of
 CC circulating HDL to circulating LDL, VLDL or total cholesterol in a human.
 CC It can also be used for decreasing the level of endogenous CETP activity
 CC in a human. The vaccine can be used to produce anti-CETP antibodies in

CC vivo and for treating cardiovascular disease.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 18; Length 21;

Best Local Similarity 100.0%; Pred. No. 4.5e-12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21

|||||

Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 5

AAW67034

ID AAW67034 standard; peptide; 21 AA.

XX

AC AAW67034;

XX

DT 15-DEC-1998 (first entry)

XX

DE Tetanus toxin fragment (residues 947-967).

XX

KW Tetanus toxin; vaccine; antibody; carbohydrate peptide conjugate;

KW dendrimeric poly-lysine; epitope; tumour.

XX

OS Clostridium tetani.

XX

PN WO9843677-A1.

XX

PD 08-OCT-1998.

XX

PF 27-MAR-1998; 98WO-EP01922.

XX

PR 27-MAR-1997; 97US-0041726.

XX

PA (INSP) INST PASTEUR.

XX

PI Bay S, Cantacuzene D, Leclerc C, Lo-man R;

XX

DR WPI; 1998-557071/47.

XX

PT Carbohydrate peptide conjugate used as vaccine - comprises carrier

PT with dendrimeric poly-lysine enabling multiple epitopes to be

PT covalently attached

XX

PS Disclosure; Page 13; 55pp; English.

XX

CC The invention relates to a new carbohydrate peptide conjugate, which

CC comprises a carrier with a dendrimeric poly-lysine enabling multiple

CC epitopes to be covalently attached to it. Also claimed are: (1) an

CC antibody purified from biological fluid or cells of organisms

CC administered with the carbohydrate peptide conjugate, and (2) a diagnosis

CC kit comprising antigen-specific antibodies elicited by immunisation with

CC the carbohydrate peptide conjugate. The peptide conjugate, antibody and

CC diagnosis kit are used to provide pharmaceutical compositions and

CC vaccines against tumours. These can be used to support an immune response

CC against viral infections caused by hepatitis virus, HIV or cytomegalo
CC virus. They can be used to enhance immune responses, especially B- and T-
CC cell responses, of humans and animals against bacterial infections. The
CC carbohydrate peptide conjugate stimulates the antibody and T-cell
CC response without stimulating undesired immune responses. The composition
CC is capable of increasing the survival of tumour bearing humans and
CC animals. The present sequence corresponds to residues 947-967 of tetanus
CC toxin. The synthetic peptide corresponding to this sequence may be used
CC as an epitope in a carbohydrate peptide conjugate.
XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 19; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
|||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 6
AAW67579

ID AAW67579 standard; peptide; 21 AA.

XX

AC AAW67579;

XX

DT 02-MAR-1999 (first entry)

XX

DE T-cell epitope peptide #5 for chimeric fimbrin/T-cell epitope peptide.

XX

KW Chimeric; non-typable Haemophilus influenzae; fimbrin; T-cell epitope;
KW immunogenic composition; immune response.

XX

OS Synthetic.

XX

PN US5843464-A.

XX

PD 01-DEC-1998.

XX

PF 02-JUN-1995; 95US-0460502.

XX

PR 02-JUN-1995; 95US-0460502.

XX

PA (OHIS) UNIV OHIO STATE.

XX

PI Bakaletz LO, Kaumaya PTP;

XX

DR WPI; 1999-044514/04.

XX

PT Synthetic chimeric fimbrin peptide - useful for vaccination against
PT non-typable Haemophilus influenzae

XX

PS Disclosure; Column 4; 16pp; English.

XX

CC The invention relates to the manufacture of a synthetic chimeric peptide
CC comprising a non-typable Haemophilus influenzae fimbrin peptide fused via

CC a linker peptide to a T-cell epitope peptide. The chimeric peptide is
CC used in immunogenic compositions which induce an immune response against
CC non-typable Haemophilus influenzae. This sequence represents an example
CC of a T-cell epitope peptide used to generate the chimeric peptide.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 20; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21

|||||

Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 7

AAW73222

ID AAW73222 standard; Protein; 21 AA.

XX

AC AAW73222;

XX

DT 25-JAN-1999 (first entry)

XX

DE Tetanus toxoid epitope.

XX

KW Multispecific single chain antibody; antibody H22; tumour cell; therapy;
KW antibody-dependent cellular cytotoxicity; ADCC; HER 2/neu; infection;
KW epidermal growth factor receptor; breast cancer; ovarian cancer.

XX

OS Synthetic.

XX

PN US5837243-A.

XX

PD 17-NOV-1998.

XX

PF 07-JUN-1996; 96US-0661052.

XX

PR 07-JUN-1996; 96US-0661052.

PR 07-JUN-1995; 95US-0484172.

XX

PA (MEDA-) MEDAREX INC.

XX

PI Deo YM, Goldstein J, Graziano R, Somasundaram C;

XX

DR WPI; 1999-023374/02.

XX

PT Specific killing of tumour cells - using a multi-specific molecule
PT comprising an anti-Fc receptor antibody and a portion which binds to
PT a target cell

XX

PS Example 7; Column 27; 57pp; English.

XX

CC This sequence represents a tetanus toxoid epitope and is recognised
CC by the multispecific single chain antibody designated H22. The
CC antibody can be used in the method of the invention for inducing
CC antibody-dependent cellular cytotoxicity (ADCC) against a tumour cell

CC which is characterised by overexpression of HER 2/neu or epidermal growth
 CC factor receptor (EGFR), comprises contacting the tumour cell with a
 CC multispecific protein molecule (preferably a single chain antibody)
 CC comprising: (a) an anti-Fc receptor antibody or an antigen binding
 CC fragment; (b) a portion which binds to HER 2/neu; and (c) a portion which
 CC binds to EGFR. The method can be used for treating cancers especially
 CC breast cancer or ovarian cancer. The multispecific antibody can also
 CC be administered prophylactically to vaccinate a subject against infection
 CC by a target cell.
 XX
 SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 20; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.5e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSWFLRVPKVSASHLE 21
 |||||
 Db 1 FNNFTVSWFLRVPKVSASHLE 21

RESULT 8
 AAB45512

ID AAB45512 standard; Protein; 21 AA..
 XX
 AC AAB45512;
 XX
 DT 26-FEB-2001 (first entry)
 XX
 DE Tetanus P30 epitope SEQ ID NO: 24.
 XX
 KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
 KW cancer; eosinophilia; vaccine; allergic rhinitis.
 XX
 OS Clostridium tetani.
 XX
 PN WO200065058-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 19-APR-2000; 2000WO-DK00205.
 XX
 PR 23-APR-1999; 99DK-0000552.
 PR 06-MAY-1999; 99US-0132811.
 XX
 PA (MEBI-) M & E BIOTECH AS.
 XX
 PI Klysner S;
 XX
 DR WPI; 2000-672791/65.
 XX
 PT Down-regulating interleukin 5 (IL-5) activity in humans by
 PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
 PT prophylaxis or amelioration of asthma or other chronic allergic
 PT conditions -
 XX
 PS Example 6; Page 137; 172pp; English.

XX
 CC The present invention is concerned with methods of treating asthma,
 CC eosinophilia, allergic rhinitis and other allergic diseases. These
 CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
 CC proteins and their coding sequences to down-regulate IL-5 activity and
 CC thus reduce eosinophil numbers. The allergic diseases may be treated
 CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
 CC it is possible that they may be used in the treatment of cancer and
 CC helminthic infections.
 XX
 SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.5e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
 |||||
 Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 9

AA99876

ID AAY99876 standard; Protein; 21 AA.

XX

AC AAY99876;

XX

DT 03-OCT-2000 (first entry)

XX

DE Tetanus toxin T cell epitope helper peptide P30.

XX

KW Human; MAGE-10; tumour rejection antigen precursor; bladder cancer;

KW prostate cancer; lung cancer; cancer detection; oesophageal cancer;

KW head and neck cancer; melanoma; myeloma; sarcoma; immunogen;

KW tetanus toxin.

XX

OS Homo sapiens.

XX

PN WO200026407-A1.

XX

PD 11-MAY-2000.

XX

PF 15-OCT-1999; 99WO-US24258.

XX

PR 30-OCT-1998; 98US-0183714.

XX

PA (LUDW-) LUDWIG INST CANCER RES.

XX

PI Boon-Falleur T, Brasseur F, Rimoldi D, Deplaen E;

XX

DR WPI; 2000-451624/39.

XX

PT Determining presence of cancer in samples, especially useful for

PT detecting bladder, prostate and lung cancer comprises assaying sample

PT for expression of tumor rejection antigen precursor MAGE-10 -

XX

PS Example 12; Page 14; 26pp; English.

XX
 CC The present sequence is a tetanus toxin T cell epitope known as
 CC Helper Peptide P30. Hybrids of this peptide and an immunogenic peptide
 CC derived from tumour rejection antigen precursor MAGE-10 were used to
 CC generate polyclonal antiserum against MAGE-10. MAGE-10 binding monoclonal
 CC antibodies can be used to detect MAGE-10 expression. A correlation
 CC between MAGE-10 expression and cancer has been discovered and thus by
 CC determining the presence of MAGE-10, the presence of cancer can be
 CC determined. MAGE-10 expression can be detected using an immunoassay, an
 CC oligonucleotide hybridisation assay or via other standard techniques.
 CC This method is especially useful for determining the presence of
 CC bladder, oesophageal, head and neck, prostate or lung cancer, or
 CC melanoma, myeloma or sarcoma.

XX
 SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.5e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
 |||||
 Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 10

AA92626

ID AA92626 standard; Protein; 21 AA.

XX

AC AA92626;

XX

DT 10-AUG-2000 (first entry)

XX

DE Foreign epitope P30.

XX

KW Foreign epitope; P2; prostate specific membrane antigen; PSM; Her2;

KW Heregulin 2; Fibroblast growth factor 8b; FGF8b; vaccination;

KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;

KW prostate cancer; cell-associated peptide antigen.

XX

OS Clostridium tetani.

XX

PN W0200020027-A2.

XX

PD 13-APR-2000.

XX

PF 05-OCT-1999; 99WO-DK00525.

XX

PR 05-OCT-1998; 98DK-0001261.

PR 20-OCT-1998; 98US-0105011.

XX

PA (MEBI-) M & E BIOTECH AS.

XX

PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;

PI Gautam A, Birk P, Karlsson G;

XX

DR WPI; 2000-349917/30.

DR N-PSDB; AAA09461.
 XX
 PT Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 XX
 PS Example 1; Page 214; 220pp; English.
 XX
 CC The claims detail a method for inducing immune responses against weakly
 CC immunogenic cell-associated peptide antigens (PA) such as those
 CC associated with cancers (i.e. self-proteins), for example, human
 CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
 CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
 CC simultaneous presentation by antigen producing cells (APCs) of the
 CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
 CC group derived from the PA and/or at least 1 B-cell group derived from the
 CC cell-associated PA; and (2) at least 1 first T helper cell group which is
 CC foreign to the animal. Analogues of human PSM, human Her2 and
 CC human/murine FGF8b comprising a substantial part of all known and
 CC predicted CTL and B-cell epitopes of the respective PA and including at
 CC least one foreign T helper epitope (e.g. P2 and/or P30) are also claimed.
 CC The method is used to treat prostate, prostate/breast or breast cancer
 CC when the PA is human PSM, FGF8b and Her2, respectively.
 XX
 SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.5e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
 |||||
 Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 11

AAAY84428

ID AAY84428 standard; Peptide; 21 AA.

XX

AC AAY84428;

XX

DT 25-JUL-2000 (first entry)

XX

DE Amino acid sequence of the tetanus toxoid P30 epitope.

XX

KW Osteoprotegerin ligand; OPGL; osteoprotegerin; osteoclastogenesis;

KW tumour necrosis factor receptor; type II transmembrane protein;

KW osteoclast differentiation; CSF-1; osteoclast activator;

KW immune response; osteoporosis; bone resorption;

KW tetanus toxoid P30 epitope.

XX

OS Clostridium tetani.

XX

PN WO200015807-A1.

XX

PD 23-MAR-2000.

XX

PF 13-SEP-1999; 99WO-DK00481.

XX
 PR 15-SEP-1998; 98DK-0001164.
 PR 02-OCT-1998; 98US-0102896.
 XX
 PA (MEBI-) M & E BIOTECH AS.
 XX
 PI Halkier T, Haaning J;
 XX
 DR WPI; 2000-271444/23.
 XX
 PT In vivo down-regulation of osteoprotegerin ligand (OPGL) activity used
 PT to treat, prevent and ameliorate osteoporosis -
 XX
 PS Example; Page 106; 110pp; English.
 XX
 CC The present sequence represents the tetanus toxoid P30 epitope. It is
 CC used to create a fusion protein with murine osteoprotegerin ligand
 CC (OPGL). Osteoprotegerin is a secreted member of the tumour necrosis
 CC factor receptor family, which blocks osteoclastogenesis in a dose
 CC dependent manner. The OPGL protein is synthesised as a type II
 CC transmembrane protein. The murine and human OPGL polypeptides are 87%
 CC homologous. OPGL is a potent osteoclast differentiation factor when
 CC combined with CSF-1. It is not capable of inducing osteoclast
 CC differentiation in the absence of CSF-1. OPGL is also an activator of
 CC mature osteoclasts. The specification describes a method for the in vivo
 CC down-regulation of OPGL activity in an animal. The method comprises
 CC using at least one OPGL polypeptide or subsequence, and/or at least one
 CC OPGL analogue to induce an immune response in the animal. The method
 CC and OPGL polypeptide are useful for treating, preventing and ameliorating
 CC osteoporosis or other diseases or conditions characterised by excessive
 CC bone resorption.
 XX
 SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.5e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
 |||||
 Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 12

AA49260

ID AA49260 standard; peptide; 21 AA.

XX

AC AA49260;

XX

DT 07-FEB-2000 (first entry)

XX

DE CD4+ T cell epitope P30TT fragment.

XX

KW Carrier protein; CD4+; T cell epitope; N6; N10; N19; immunogen;

KW encapsulated bacteria.

XX

OS Synthetic.

XX
 PN WO9955730-A2.
 XX
 PD 04-NOV-1999.
 XX
 PF 27-APR-1999; 99WO-IB00844.
 XX
 PR 27-APR-1998; 98GB-0008932.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Rappuoli R, Grandi G;
 XX
 DR WPI; 2000-023325/02.
 XX
 PT Carrier proteins containing CD4+ epitopes useful for protecting against
 PT diseases caused by encapsulated bacteria -
 XX
 PS Disclosure; Page 36; 76pp; English.
 XX
 CC The invention provides carrier proteins comprising at least 5 CD4+ T
 CC cell epitope. The carrier protein comprises at least 1 of N6, N10 or
 CC N19. The carrier protein can be prepared by expressing a vector
 CC comprising a nucleic acid molecule encoding the protein, in a host cell
 CC and recovering the expressed protein. The carrier protein can also be
 CC produced by (a) constructing oligonucleotide molecules that encode
 CC peptide epitopes; (b) annealing the oligonucleotides to form duplexes;
 CC (c) introducing the duplexes into an expression vector; (d) introducing
 CC the expression vector into a host cell; and (e) isolating the fusion
 CC protein produced from a culture of the host cells. The carrier protein
 CC can be used as a protective immunogen in the control of diseases caused
 CC by encapsulated bacteria. Sequences AAY49256-266 represent CD4+ T cell
 CC epitopes inserted in the recombinant polyepitope carrier proteins.
 XX
 SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.5e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
 |||||
 Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 13

AAE11764

ID AAE11764 standard; peptide; 21 AA.

XX

AC AAE11764;

XX

DT 18-DEC-2001 (first entry)

XX

DE Clostridium tetani P30 epitope.

XX

KW Amyloid protein; neuroprotective; nootropic; immunostimulant; vaccine;
 KW Alzheimer's disease; anticonvulsant; gene therapy; Pick's disease;

RESULT 14

AAB85702

ID AAB85702 standard; peptide; 21 AA.

XX

AC AAB85702;

XX

DT 29-OCT-2001 (first entry)

XX

DE Amino acid sequence of P30 epitope.

XX

KW Multivalent protein; immune response; Plasmodium vivax; parasite;
KW protozoacide; vaccine; malaria; recombinant; ViVac1; ViVac2.

XX

OS Plasmodium vivax.

XX

PN WO200155181-A2.

XX

PD 02-AUG-2001.

XX

PF 29-JAN-2001; 2001WO-US02937.

XX

PR 31-JAN-2000; 2000US-0179213.

XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Lal AA, Xiao L, Zhou Z;

XX

DR WPI; 2001-514557/56.

XX

PT New recombinant multivalent protein comprising antigenic determinants
PT derived from more than one stage in a life cycle of Plasmodium vivax,
PT useful as a vaccine for treating, preventing and reducing malarial
PT infection -

XX

PS Example 1; Page 25; 59pp; English.

XX

CC The invention relates to recombinant multivalent proteins (I) that
CC stimulate an immune response to Plasmodium vivax. (I) comprises antigenic
CC determinants, fragments or conservative substitutions, derived from more
CC than one stage in a life cycle of a Plasmodium vivax parasite. (I) is
CC useful as a vaccine for stimulating an immune response, specifically a
CC protective immune response that confers increased resistance to infection
CC by Plasmodium parasites, such as P. vivax. (I) is especially useful in
CC the treatment, prevention and reduction of malarial infection, as
CC research or diagnostic reagents for the detection of Plasmodium species
CC in a biological sample, and for conferring immunity against multiple
CC stages of the malarial parasite. The antibodies produced are useful for
CC the detection or measurement of antigenic epitopes derived from one or
CC more stages in a life cycle of a parasite, particularly P. vivax. The
CC vaccine comprising the recombinant proteins, is cost-effective, health-
CC promoting intervention for controlling, preventing or treating the
CC incidence of malaria. The present sequence represents the amino acid
CC sequence of a P30 epitope, a component of the multivalent and
CC multistage proteins ViVac1p and ViVac2p.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 15

AAB85453

ID AAB85453 standard; peptide; 21 AA.

XX

AC AAB85453;

XX

DT 25-SEP-2001 (first entry)

XX

DE Universal tetanus toxin Th epitope TT947-967.

XX

KW HER 2/neu; epidermal growth factor receptor; EGFR; multispecific protein;
KW Fc receptor; FcR; tumor cell; breast; cancer; sarcoma; carcinoma; HIV;
KW pathogenic; Toxoplasma gondii; candidiasis; systemic lupus; cytostatic;
KW immune thrombocytopenia purpura; immunosuppressive; antiviral;
KW antifungal; antiprotozoal; tetanus toxin.

XX

OS Clostridium tetani.

XX

PN US6270765-B1.

XX

PD 07-AUG-2001.

XX

PF 06-NOV-1998; 98US-0188082.

XX

PR 07-JUN-1996; 96US-0661052.

PR 07-JUN-1995; 95US-0484172.

XX

PA (MEDA-) MEDAREX INC.

XX

PI Deo YM, Goldstein J, Graziano R, Somasundaram C;

XX

DR WPI; 2001-475189/51.

XX

PT Inducing killing of tumor cells which expresses HER 2/neu or epidermal
PT growth factor receptor (EGFR) by contacting the cell with multispecific
PT proteins comprising an anti-Fc receptor, -Her 2/neu or -EGFR antibody,
PT useful for treating cancer -

XX

PS Example 7; Column 29; 57pp; English.

XX

CC The invention relates to a new method for inducing killing of a tumor
CC cell which expresses HER 2/neu or epidermal growth factor receptor
CC (EGFR). The method comprises contacting the tumor cell with a
CC multispecific protein comprising a component, preferably an antibody,
CC which binds to an Fc receptor (FcR), Her 2/neu or EGFR. The method is
CC useful for inducing killing of a tumor cell from breast cancer, sarcoma,
CC carcinoma, or ovarian cancer. Specific multispecific proteins can also be

CC administered to a subject to treat or prevent other diseases or
CC conditions, including pathogenic infections (e.g., viral (such as HIV)),
CC protozoan infections (such as Toxoplasma gondii), fungal infections
CC (such as candidiasis), and an autoimmunity (e.g. immune thrombocytopenia
CC purpura and systemic lupus). The present sequence represents an universal
CC tetanus toxin Th epitope TT947-967.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 112; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

Search completed: December 13, 2003, 13:34:23
Job time : 3.48718 secs

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 1.28205 Seconds
 (without alignments)
 693.052 Million cell updates/sec

Title: US-09-785-215-6
 Perfect score: 112
 Sequence: 1 FNNFTVSFWLRVPKVSASHLE 21

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued_Patents_AA:*
 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	112	100.0	21	1	US-07-610-525-1	Sequence 1, Appli
2	112	100.0	21	2	US-08-661-052-12	Sequence 12, Appl
3	112	100.0	21	2	US-08-460-502-8	Sequence 8, Appli
4	112	100.0	21	2	US-08-724-774B-5	Sequence 5, Appli
5	112	100.0	21	3	US-09-089-595-5	Sequence 5, Appli
6	112	100.0	21	3	US-09-382-855-5	Sequence 5, Appli
7	112	100.0	21	3	US-09-183-714B-5	Sequence 5, Appli
8	112	100.0	21	3	US-09-188-082-12	Sequence 12, Appl
9	112	100.0	21	3	US-09-171-969-10	Sequence 10, Appl
10	112	100.0	21	4	US-09-364-088-12	Sequence 12, Appl
11	112	100.0	21	4	US-09-642-281-5	Sequence 5, Appli

12	112	100.0	21	4	US-09-102-716-12	Sequence 12, Appl
13	112	100.0	21	4	US-08-432-483A-3	Sequence 3, Appli
14	112	100.0	21	4	US-09-148-711A-8	Sequence 8, Appli
15	112	100.0	21	4	US-09-589-717-5	Sequence 5, Appli
16	112	100.0	21	4	US-08-945-289-3	Sequence 3, Appli
17	112	100.0	21	5	PCT-US93-11703-66	Sequence 66, Appl
18	112	100.0	31	5	PCT-US93-11703-64	Sequence 64, Appl
19	112	100.0	32	1	US-08-446-692-14	Sequence 14, Appl
20	112	100.0	32	2	US-08-488-351A-14	Sequence 14, Appl
21	112	100.0	452	1	US-07-618-312A-2	Sequence 2, Appli
22	112	100.0	452	1	US-07-618-312A-4	Sequence 4, Appli
23	112	100.0	452	1	US-08-110-786A-8	Sequence 8, Appli
24	112	100.0	452	1	US-08-280-228-2	Sequence 2, Appli
25	112	100.0	452	1	US-08-280-228-4	Sequence 4, Appli
26	112	100.0	618	1	US-08-668-381A-5	Sequence 5, Appli
27	112	100.0	853	4	US-08-913-880C-17	Sequence 17, Appl
28	112	100.0	858	4	US-08-913-880C-16	Sequence 16, Appl
29	112	100.0	860	4	US-08-913-880C-15	Sequence 15, Appl
30	112	100.0	862	4	US-08-913-880C-14	Sequence 14, Appl
31	112	100.0	865	4	US-08-913-880C-13	Sequence 13, Appl
32	112	100.0	866	4	US-08-913-880C-12	Sequence 12, Appl
33	112	100.0	874	4	US-08-913-880C-11	Sequence 11, Appl
34	112	100.0	875	4	US-08-913-880C-10	Sequence 10, Appl
35	112	100.0	1315	4	US-08-913-880C-1	Sequence 1, Appli
36	107	95.5	22	1	US-08-446-692-5	Sequence 5, Appli
37	107	95.5	22	2	US-08-488-351A-5	Sequence 5, Appli
38	107	95.5	22	3	US-09-100-409A-41	Sequence 41, Appl
39	107	95.5	22	5	PCT-US95-13841-8	Sequence 8, Appli
40	100	89.3	19	1	US-07-610-525-2	Sequence 2, Appli
41	94.5	84.4	20	2	US-08-319-704-11	Sequence 11, Appl
42	88	78.6	17	1	US-07-610-525-3	Sequence 3, Appli
43	87.5	78.1	22	2	US-08-817-933A-8	Sequence 8, Appli
44	79	70.5	14	1	US-07-610-525-6	Sequence 6, Appli
45	79	70.5	14	1	US-08-787-547-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1

US-07-610-525-1

; Sequence 1, Application US/07610525

; Patent No. 5196512

; GENERAL INFORMATION:

; APPLICANT: BIANCHI Eisabetta

; APPLICANT: PESSI Antonello

; APPLICANT: CORRADIN Giampietro

; TITLE OF INVENTION: SYNTHETIC PEPTIDES USEFUL AS

; TITLE OF INVENTION: UNIVERSAL CARRIERS FOR THE PREPARATION OF IMMUNOGENIC CONJUGATES

; TITLE OF INVENTION: AND THEIR USE IN THE DEVELOPMENT OF SYNTHETIC VACCINES.

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SHEA & GOULD

; STREET: 1251 AVENUE OF THE AMERICAS

; CITY: NEW YORK

```

; STATE: NEW YORK
; COUNTRY: UNITED STATES
; ZIP: 10020-1193
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/610,525
; FILING DATE: 19901108
; CLASSIFICATION: 424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-827-3000
; TELEFAX: 212-840-6702
; TELEX: 423973
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acid residues
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
; FRAGMENT TYPE: internal fragment
US-07-610-525-1

```

```

Query Match          100.0%; Score 112; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
        ||||||||||||||||
Db      1 FNNFTVSFWLRVPKVSASHLE 21

```

RESULT 2

```

US-08-661-052-12
; Sequence 12, Application US/08661052
; Patent No. 5837243
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-Fc RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```



```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/661,052
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-661-052-12

```

```

Query Match          100.0%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
        |||||||||||||||||||||
Db      1 FNNFTVSFWLRVPKVSASHLE 21

```

RESULT 3

```

US-08-460-502-8
; Sequence 8, Application US/08460502
; Patent No. 5843464
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Parvin T.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,502

```

```

; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Golrick, Mary E.
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-502-8

```

```

Query Match          100.0%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
        |||||
Db      1 FNNFTVSFWLRVPKVSASHLE 21

```

RESULT 4

US-08-724-774B-5

```

; Sequence 5, Application US/08724774B
; Patent No. 5908778
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie,
; APPLICANT: Pierre; Cerrottini, Jean-Charles; Carrel,
; APPLICANT: Stefan; Reed, Daryl
; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
; TITLE OF INVENTION: Rejection Antigen Precuros Mage-10,
; TITLE OF INVENTION: Antibodies Specific To The Molecule, and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,774B
; FILING DATE: 03-October-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Hanson, No. 5908778man D.
 ; REGISTRATION NUMBER: 30,946
 ; REFERENCE/DOCKET NUMBER: LUD 5457
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 688-9200
 ; TELEFAX: (212) 838-3884
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 21
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 US-08-724-774B-5

Query Match 100.0%; Score 112; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.7e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSWFLRVPKVSASHLE 21
 |||||
 Db 1 FNNFTVSWFLRVPKVSASHLE 21

RESULT 5

US-09-089-595-5

; Sequence 5, Application US/09089595

; Patent No. 6153728

; GENERAL INFORMATION:

; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie,
 ; APPLICANT: Pierre; Cerrottini, Jean-Charles; Carrel,
 ; APPLICANT: Stefan; Reed, Daryl
 ; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
 ; TITLE OF INVENTION: Rejection Antigen Precuros Mage-10,
 ; TITLE OF INVENTION: Antibodies Specific To The Molecule, and
 ; TITLE OF INVENTION: Uses Thereof
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Felfe & Lynch
 ; STREET: 805 Third Avenue
 ; CITY: New York City
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
 ; COMPUTER: IBM
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/089,595
 ; FILING DATE:
 ; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/724,774
 ; FILING DATE: 03-October-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 6153728man D.
 ; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 5457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
US-09-089-595-5

Query Match 100.0%; Score 112; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSPWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSPWLRVPKVSASHLE 21

RESULT 6

US-09-382-855-5

; Sequence 5, Application US/09382855
; Patent No. 6174692

; GENERAL INFORMATION:

; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie, Pierre;
; APPLICANT: Cerrottini, Jean-Charles; Carrel, Stefan; Reed, Daryl
; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor Rejection
; TITLE OF INVENTION: Antigen Precursors Mage-10, Antibodies Specific To The
Molecule, a

; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/382,855
; FILING DATE: 25-August-1999

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 09/089,595
; FILING DATE: 02-June-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/724,774
; FILING DATE: 03-October-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Schofield, Mary Anne
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5457.2 DIV - JEL/MAS

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
US-09-382-855-5
```

```
Query Match          100.0%; Score 112; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 FNNFTVSFWLRVPKVSASHLE 21
          ||||||||||||||||
Db      1 FNNFTVSFWLRVPKVSASHLE 21
```

RESULT 7

```
US-09-183-714B-5
; Sequence 5, Application US/09183714B
; Patent No. 6221593
; GENERAL INFORMATION:
; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: Brasseur, Francis
; APPLICANT: Rimoldi, Donata
; APPLICANT: De Plaen, Etienne
; TITLE OF INVENTION: Method for Determining Cancer by Determining Expression
; TITLE OF INVENTION: of MAGE-10
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/183,714B
; CURRENT FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 08/724,774
; PRIOR FILING DATE: 1996-10-03
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-183-714B-5
```

```
Query Match          100.0%; Score 112; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 FNNFTVSFWLRVPKVSASHLE 21
          ||||||||||||||||
Db      1 FNNFTVSFWLRVPKVSASHLE 21
```

RESULT 8

```
US-09-188-082-12
; Sequence 12, Application US/09188082
; Patent No. 6270765
; GENERAL INFORMATION:
```

```

; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-Fc RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,082
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/661,052
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-188-082-12

```

```

Query Match          100.0%; Score 112; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
        ||||||||||||||||
Db      1 FNNFTVSFWLRVPKVSASHLE 21

```

RESULT 9

US-09-171-969-10

; Sequence 10, Application US/09171969

; Patent No. 6284533

; GENERAL INFORMATION:

; APPLICANT: Thomas, Lawrence J.

; TITLE OF INVENTION: PLASMID-BASED VACCINE FOR TREATING ATHEROSCLEROSIS

```

; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 75 State Street, Suite 2300
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1807
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,969
; FILING DATE: 01 May 1997 (01.05.97)
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/640,713
; FILING DATE: 01 May 1996 (01.05.96)
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/802,967
; FILING DATE: 21 February 1997 (21.02.97)
; ATTORNEY/AGENT INFORMATION:
; NAME: Leon R. Yankwich
; REGISTRATION NUMBER: 30,237
; REFERENCE/DOCKET NUMBER: TCS 414.1 PCT (05872)
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
; ANTI-SENSE:
; FEATURE:
; NAME/KEY:
; LOCATION:
US-09-171-969-10

```

```

Query Match          100.0%; Score 112; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
        ||||||||||||||||
Db      1 FNNFTVSFWLRVPKVSASHLE 21

```

```

RESULT 10
US-09-364-088-12
; Sequence 12, Application US/09364088
; Patent No. 6365161
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo, et al.
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-Fc RECEPTOR ANTIBODIES

```

```

; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street, 24th Floor
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/364,088
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/188,082
; FILING DATE: 07-JUNE-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remillard, Jane E.
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: MXI-043CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-7414
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-364-088-12

```

```

Query Match          100.0%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
        ||||||||||||||||
Db      1 FNNFTVSFWLRVPKVSASHLE 21

```

RESULT 11

US-09-642-281-5

; Sequence 5, Application US/09642281

; Patent No. 6387698

; GENERAL INFORMATION:

; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie, Pierre;

; Cerrottini, Jean-Charles; Carrel, Stefan; Reed, Daryl

; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor Rejection

; Antigen Precursors Mage-10, Antibodies Specific To
The Molecule, and Uses Thereof


```

;      NUMBER OF SEQUENCES: 16
;      CORRESPONDENCE ADDRESS:
;          ADDRESSEE: LAHIVE & COCKFIELD
;          STREET: 60 State Street, Suite 510
;          CITY: Boston
;          STATE: Massachusetts
;          COUNTRY: USA
;          ZIP: 02109-1875
;      COMPUTER READABLE FORM:
;          MEDIUM TYPE: Floppy disk
;          COMPUTER: IBM PC compatible
;          OPERATING SYSTEM: PC-DOS/MS-DOS
;          SOFTWARE: PatentIn Release #1.0, Version #1.25
;      CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: US/09/102,716
;          FILING DATE: 22-Jun-1998
;      PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: 08/661,052
;          FILING DATE: <Unknown>
;      ATTORNEY/AGENT INFORMATION:
;          NAME: Arnold, Beth E.
;          REGISTRATION NUMBER: 35,430
;          REFERENCE/DOCKET NUMBER: MXI-043CP
;      TELECOMMUNICATION INFORMATION:
;          TELEPHONE: (617)227-7400
;          TELEFAX: (617)227-5941
;      INFORMATION FOR SEQ ID NO: 12:
;          SEQUENCE CHARACTERISTICS:
;              LENGTH: 21 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;          MOLECULE TYPE: peptide
;          FRAGMENT TYPE: internal
;          SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-102-716-12

```

```

Query Match          100.0%;  Score 112;  DB 4;  Length 21;
Best Local Similarity 100.0%;  Pred. No. 1.7e-12;
Matches 21;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
        |||||||||||||||||||||
Db      1 FNNFTVSFWLRVPKVSASHLE 21

```

```

RESULT 13
US-08-432-483A-3
; Sequence 3, Application US/08432483A
; Patent No. 6410022
; GENERAL INFORMATION:
;     APPLICANT: Rittershaus, Charles W.
;     TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER
;     TITLE OF INVENTION: TRANSFER PROTEIN (CETP) ACTIVITY
;     NUMBER OF SEQUENCES: 5
;     CORRESPONDENCE ADDRESS:
;         ADDRESSEE: Banner & Witcoff, Ltd.
;         STREET: Ten South Wacker Drive

```

```

; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-7407
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,483A
; FILING DATE: 1-May-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Leon R. Yankwich
; REGISTRATION NUMBER: 30,237
; REFERENCE/DOCKET NUMBER: 95,179 (TCS-95179)
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL:
; ANTI-SENSE:
; FEATURE:
; NAME/KEY: 21-amino acid tetanus toxoid universal
; NAME/KEY: helper T cell epitope.
; LOCATION:
; PUBLICATION INFORMATION:
; AUTHORS: Panina-Bordignon, P., et al.
; TITLE: Universally immunogenic T cell
; TITLE: epitopes: promiscuous binding to human MHC class II and
; TITLE: promiscuous recognition by T cells
; JOURNAL: European Journal of Immunology
; VOLUME: 19
; ISSUE:
; PAGES: 2237-2242
; DATE: 1989
; RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 21
US-08-432-483A-3

```

```

Query Match          100.0%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
        ||||||||||||||||
Db      1 FNNFTVSFWLRVPKVSASHLE 21

```

```

RESULT 14
US-09-148-711A-8
; Sequence 8, Application US/09148711A

```

; Patent No. 6436405
; GENERAL INFORMATION:
; APPLICANT: The Ohio State University
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; FILE REFERENCE: 18525-04010
; CURRENT APPLICATION NUMBER: US/09/148,711A
; CURRENT FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 08/460,502
; PRIOR FILING DATE: 1995-06-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 21
; TYPE: PRT
; ORGANISM: clostridium tetani
US-09-148-711A-8

Query Match 100.0%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
|||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 15

US-09-589-717-5

; Sequence 5, Application US/09589717
; Patent No. 6497879
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie,
; Pierre; Cerrottini, Jean-Charles; Carrel,
; Stefan; Reed, Daryl
; TITLE OF INVENTION: MAGe-10 ENCODING cDNA, The Tumor
; Rejection Antigen Precuros Mage-10,
; Antibodies Specific To The Molecule, and
; Uses Thereof
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/589,717
; FILING DATE: 08-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/089,595

; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6497879man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-589-717-5

Query Match 100.0%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
|||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

Search completed: December 13, 2003, 13:31:54
Job time : 2.28205 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:30:00 ; Search time 2.38462 Seconds
(without alignments)
1637.856 Million cell updates/sec

Title: US-09-785-215-6
Perfect score: 112
Sequence: 1 FNNFTVSFWLRVPKVSASHLE 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Match	Length	DB	ID	Description
		%				
		Query				

1	112	100.0	21	9	US-09-943-548-3	Sequence 3, Appli
2	112	100.0	21	9	US-09-848-834A-4	Sequence 4, Appli
3	112	100.0	21	10	US-09-785-215-6	Sequence 6, Appli
4	112	100.0	21	12	US-09-405-986-2	Sequence 2, Appli
5	112	100.0	21	12	US-10-223-809A-6	Sequence 6, Appli
6	112	100.0	21	12	US-10-261-208-5	Sequence 5, Appli
7	112	100.0	21	12	US-10-295-074-5	Sequence 5, Appli
8	112	100.0	21	15	US-10-204-362-6	Sequence 6, Appli
9	112	100.0	21	15	US-10-339-522-3	Sequence 3, Appli
10	112	100.0	21	15	US-10-223-711-8	Sequence 8, Appli
11	112	100.0	34	9	US-09-848-834A-10	Sequence 10, Appl
12	112	100.0	37	9	US-09-848-834A-14	Sequence 14, Appl
13	112	100.0	50	9	US-09-848-834A-18	Sequence 18, Appl
14	112	100.0	158	12	US-10-297-942-2	Sequence 2, Appli
15	112	100.0	158	12	US-10-297-942-10	Sequence 10, Appl
16	112	100.0	158	12	US-10-297-942-12	Sequence 12, Appl
17	112	100.0	158	12	US-10-297-942-18	Sequence 18, Appl
18	112	100.0	158	12	US-10-297-942-20	Sequence 20, Appl
19	112	100.0	194	12	US-10-295-074-46	Sequence 46, Appl
20	112	100.0	194	12	US-10-295-074-47	Sequence 47, Appl
21	112	100.0	285	12	US-10-295-074-9	Sequence 9, Appli
22	112	100.0	285	12	US-10-295-074-11	Sequence 11, Appl
23	112	100.0	287	12	US-10-295-074-13	Sequence 13, Appl
24	112	100.0	287	12	US-10-295-074-15	Sequence 15, Appl
25	112	100.0	463	11	US-09-816-467-2	Sequence 2, Appli
26	112	100.0	514	12	US-10-295-074-49	Sequence 49, Appl
27	112	100.0	514	12	US-10-295-074-51	Sequence 51, Appl
28	112	100.0	514	12	US-10-295-074-59	Sequence 59, Appl
29	112	100.0	517	12	US-10-295-074-53	Sequence 53, Appl
30	112	100.0	605	12	US-10-130-973A-11	Sequence 11, Appl
31	112	100.0	665	12	US-10-130-973A-9	Sequence 9, Appli
32	112	100.0	882	12	US-10-130-973A-3	Sequence 3, Appli
33	112	100.0	907	12	US-10-130-973A-5	Sequence 5, Appli
34	112	100.0	1052	12	US-10-130-973A-17	Sequence 17, Appl
35	112	100.0	1112	12	US-10-130-973A-16	Sequence 16, Appl
36	112	100.0	1315	12	US-10-241-596-141	Sequence 141, App
37	107	95.5	22	11	US-09-865-294-4	Sequence 4, Appli
38	95.5	85.3	20	14	US-10-044-034-20	Sequence 20, Appl
39	88	78.6	22	11	US-09-747-802-12	Sequence 12, Appl
40	79	70.5	14	10	US-09-909-460-43	Sequence 43, Appl
41	79	70.5	15	10	US-09-909-460-44	Sequence 44, Appl
42	61	54.5	144	8	US-08-981-087A-2	Sequence 2, Appli
43	61	54.5	431	8	US-08-981-087A-1	Sequence 1, Appli
44	61	54.5	432	11	US-09-910-186A-16	Sequence 16, Appl
45	61	54.5	432	11	US-09-910-186A-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
 US-09-943-548-3
 ; Sequence 3, Application US/09943548
 ; Patent No. US20020042364A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rittershaus, Charles W.

```

; APPLICANT: Thomas, Lawrence J.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP)
ACTIVITY
; FILE REFERENCE: TCS-411.1P US-1; Tcs-411.1P US-2
; CURRENT APPLICATION NUMBER: US/09/943,548
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: helper T cell epitope of tetanus toxin
US-09-943-548-3

```

```

Query Match          100.0%; Score 112; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
        |||||||||||||||||||
Db      1 FNNFTVSFWLRVPKVSASHLE 21

```

```

RESULT 2
US-09-848-834A-4
; Sequence 4, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Tetanus bacillus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(21)
; OTHER INFORMATION: Amino acid sequence 947-967 of Tetanus
; OTHER INFORMATION: Toxoid Precursor (Tentoxylisin)
US-09-848-834A-4

```

```

Query Match          100.0%; Score 112; DB 9; Length 21;

```


Best Local Similarity 100.0%; Pred. No. 7.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 3

US-09-785-215-6

; Sequence 6, Application US/09785215
; Publication No. US20020187157A1
; GENERAL INFORMATION:
; APPLICANT: JENSEN, Martin Roland et al.
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID
; FILE REFERENCE: 3631-0107P
; CURRENT APPLICATION NUMBER: US/09/785,215
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-785-215-6

Query Match 100.0%; Score 112; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 4

US-09-405-986-2

; Sequence 2, Application US/09405986
; Publication No. US20030157115A1
; GENERAL INFORMATION:
; APPLICANT: BAY, Sylvie
; APPLICANT: CANTACUZENE, Daniele
; APPLICANT: LECLERC, Claude
; APPLICANT: LO-MAN, Richard
; TITLE OF INVENTION: MULTIPLE ANTIGEN GLYCOPEPTIDE CARBOHYDRATE, VACCINE
; TITLE OF INVENTION: COMPRISING THE SAME AND USE THEREOF
; FILE REFERENCE: 134I US 3565
; CURRENT APPLICATION NUMBER: US/09/405,986
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: 60/041,726
; EARLIER FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani

US-09-405-986-2

Query Match 100.0%; Score 112; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 5

US-10-223-809A-6

; Sequence 6, Application US/10223809A
; Publication No. US20030157117A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; APPLICANT: Rasmussen, Peter Birk et al.
; TITLE OF INVENTION: No. US20030157117A1el Method for Down-Regulation of Amyloid
; FILE REFERENCE: 674542-2008
; CURRENT APPLICATION NUMBER: US/10/223,809A
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: US 60/337,543
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: US 60/373,027
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: DE 2001 01231
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: DE 2002 0058
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-223-809A-6

Query Match 100.0%; Score 112; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 6

US-10-261-208-5

; Sequence 5, Application US/10261208
; Publication No. US20030158388A1
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie,
; Pierre; Cerrottini, Jean-Charles; Carrel,
; Stefan; Reed, Daryl
; TITLE OF INVENTION: MAG-10 ENCODING cDNA, The Tumor

; Rejection Antigen Precuros Mage-10,
; Antibodies Specific To The Molecule, and
; Uses Thereof

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/261,208

; FILING DATE: 30-Sep-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/589,717

; FILING DATE: 08-Jun-2000

; APPLICATION NUMBER: 09/089,595

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. US20030158388A1man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 5457

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 21

; TYPE: amino acid

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-10-261-208-5

Query Match 100.0%; Score 112; DB 12; Length 21;

Best Local Similarity 100.0%; Pred. No. 7.1e-11;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21

|||||

Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 7

US-10-295-074-5

; Sequence 5, Application US/10295074

; Publication No. US20030185845A1

; GENERAL INFORMATION:

; APPLICANT: Pharmexa A/S

; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS

; FILE REFERENCE: P1013DK00

; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Tetanus toxoid P30 epitope
US-10-295-074-5

Query Match 100.0%; Score 112; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 8

US-10-204-362-6
; Sequence 6, Application US/10204362
; Publication No. US20030086938A1
; GENERAL INFORMATION:
; APPLICANT: M&E Biotech A/S
; TITLE OF INVENTION: No. US20030086938A1el Method For Down-Regulation Of Amyloid
; FILE REFERENCE: 3631-0120P
; CURRENT APPLICATION NUMBER: US/10/204,362
; CURRENT FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-204-362-6

Query Match 100.0%; Score 112; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 9

US-10-339-522-3
; Sequence 3, Application US/10339522
; Publication No. US20030108559A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; APPLICANT: Thomas, Lawrence J.

; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP)
ACTIVITY

; FILE REFERENCE: TCS-411.1P US-3
; CURRENT APPLICATION NUMBER: US/10/339,522
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 09/943,334
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/943,548
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: helper T cell epitope of tetanus toxin
US-10-339-522-3

Query Match 100.0%; Score 112; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
|||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 10

US-10-223-711-8
; Sequence 8, Application US/10223711
; Publication No. US20030113344A1
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Pravin T.P.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; FILE REFERENCE: 18525/04058
; CURRENT APPLICATION NUMBER: US/10/223,711
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 09/148,711
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 08/460,502
; PRIOR FILING DATE: 1995-06-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-223-711-8

Query Match 100.0%; Score 112; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 11

US-09-848-834A-10
; Sequence 10, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 947-967 of the
; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to am
; OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated phenylalanine
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(21)
; OTHER INFORMATION: Amino acids 947-967 of the Tetanus Toxoid Precursor
; OTHER INFORMATION: (Tentoxylisin)
; NAME/KEY: PEPTIDE
; LOCATION: (22)..(25)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (26)..(34)
; OTHER INFORMATION: Amino acids 2-10 of the human GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (34)..(34)
; OTHER INFORMATION: Amidated glycine or glycinamide
US-09-848-834A-10

Query Match 100.0%; Score 112; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | |

Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 12

US-09-848-834A-14

; Sequence 14, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10
of the G
; OTHER INFORMATION: nRH hormone linked by a spacer to amino acid sequence
947-967 of
; OTHER INFORMATION: the Tetanus toxoid precursor (Tentoxylisin)
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(37)
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid
precursor
; OTHER INFORMATION: (Tentoxylisin)
US-09-848-834A-14

Query Match 100.0%; Score 112; DB 9; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 17 FNNFTVSFWLRVPKVSASHLE 37

RESULT 13

US-09-848-834A-18

; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:

```

; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10
of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of
the Tet
; OTHER INFORMATION: anus toxoid precursor (Tentoxylisin) protein linked by a
spacer t
; OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (50)..(50)
; OTHER INFORMATION: Amidated glycine or glycinamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(37)
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid
precursor (Tent
; OTHER INFORMATION: oxylysin
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(41)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (42)..(50)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18

```

```

Query Match          100.0%; Score 112; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
        ||||||||||||||||
Db      17 FNNFTVSFWLRVPKVSASHLE 37

```

```

RESULT 14
US-10-297-942-2

```



```

; Sequence 2, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
; APPLICANT: Ferring BV
; TITLE OF INVENTION: Solubilised Protein Vaccines
; FILE REFERENCE: P68445US0
; CURRENT APPLICATION NUMBER: US/10/297,942
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: PCT/DK01/00431
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: DK PA 2000 00966
; PRIOR FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-297-942-2

```

```

Query Match          100.0%; Score 112; DB 12; Length 158;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
          ||||||||||||||||
Db      65 FNNFTVSFWLRVPKVSASHLE 85

```

RESULT 15

```

US-10-297-942-10
; Sequence 10, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
; APPLICANT: Ferring BV
; TITLE OF INVENTION: Solubilised Protein Vaccines
; FILE REFERENCE: P68445US0
; CURRENT APPLICATION NUMBER: US/10/297,942
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: PCT/DK01/00431
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: DK PA 2000 00966
; PRIOR FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-297-942-10

```

```

Query Match          100.0%; Score 112; DB 12; Length 158;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FNNFTVSFWLRVPKVSASHLE 21
          ||||||||||||||||

```

Db

41 FNNFTVSEFWLRVPKVSASHLE 61

Search completed: December 13, 2003, 13:36:50

Job time : 3.38462 secs

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 1.30769 Seconds
(without alignments)
1544.356 Million cell updates/sec

Title: US-09-785-215-6
Perfect score: 112
Sequence: 1 FNNFTVSFWLRVPKVSASHLE 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	112	100.0	1315	1	BTCLTN	tentoxilysin (EC 3
2	62	55.4	1268	2	S33411	botulinum neurotox
3	61	54.5	366	2	S48110	neurotoxin type F
4	61	54.5	369	2	S48109	neurotoxin type F
5	61	54.5	1274	2	I40813	neurotoxin type F
6	61	54.5	1297	2	S39791	neurotoxin - Clost
7	59	52.7	1296	1	BTCLAB	bontoxilysin (EC 3
8	58	51.8	1291	1	A48940	bontoxilysin (EC 3
9	58	51.8	1291	2	I40631	non-proteolytic bo
10	56	50.0	367	2	S48106	neurotoxin type E
11	56	50.0	1251	2	JH0256	botulinum neurotox
12	56	50.0	1252	2	S21178	botulinum neurotox
13	56	50.0	1296	2	I40645	botulinum neurotox

14	52	46.4	449	2	S23158	nucleocapsid prote
15	52	46.4	464	1	MNVUWC	nonstructural prot
16	52	46.4	467	1	MNVUW1	nonstructural prot
17	51	45.5	1196	2	JQ1467	toxin, nontoxic co
18	51	45.5	1196	2	S46430	botulinum neurotox
19	49	43.8	276	2	T33493	hypothetical prote
20	48	42.9	504	2	T47446	hypothetical prote
21	48	42.9	1285	2	S70582	botulinum neurotox
22	48	42.9	1291	2	S46431	botulinum neurotox
23	48	42.9	1291	2	A49777	botulinum neurotox
24	47.5	42.4	1276	2	S11455	botulinum neurotox
25	47	42.0	359	2	F87937	protein F14B6.6 [i
26	47	42.0	385	2	T20879	hypothetical prote
27	47	42.0	469	2	B37837	probable alpha-amy
28	46	41.1	322	2	T25966	hypothetical prote
29	46	41.1	442	2	I47074	gene CD5 protein -
30	46	41.1	496	2	T38197	probable myb-like
31	46	41.1	753	2	C96668	unknown protein F1
32	46	41.1	1162	2	A47708	progenitor toxin n
33	46	41.1	1162	2	I40817	botulinum toxin no
34	45	40.2	528	2	T41362	hypothetical prote
35	45	40.2	886	2	T39081	hypothetical prote
36	44	39.3	209	2	A82470	conserved hypothet
37	44	39.3	404	2	AD0574	enterochelin ester
38	44	39.3	591	2	S04401	spheroidene monoox
39	43	38.4	152	2	D64943	probable membrane
40	43	38.4	152	2	B99945	hypothetical prote
41	43	38.4	152	2	E85793	hypothetical prote
42	43	38.4	381	2	A90711	hypothetical prote
43	43	38.4	381	2	E85561	hypothetical prote
44	43	38.4	514	2	T20858	hypothetical prote
45	43	38.4	781	2	F83884	hypothetical prote

ALIGNMENTS

RESULT 1

BTCLTN

tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani

N;Alternate names: tetanus neurotoxin

C;Species: Clostridium tetani

C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 03-Jun-2002

C;Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364

R;Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, M.; Habermann, E.; Niemann, H.

EMBO J. 5, 2495-2502, 1986

A;Title: Tetanus toxin: primary structure, expression in E. coli, and homology with botulinum toxins.

A;Reference number: A25689; MUID:87053814; PMID:3536478

A;Accession: A25689

A;Molecule type: DNA

A;Residues: 1-1315 <EIS>

A;Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770

R;Fairweather, N.F.; Lyness, V.A.

Nucleic Acids Res. 14, 7809-7812, 1986

A;Title: The complete nucleotide sequence of tetanus toxin.

A;Reference number: A25757; MUID:87040747; PMID:3774547
 A;Accession: A25757
 A;Molecule type: DNA
 A;Residues: 1-1315 <FAI>
 A;Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774
 A;Experimental source: strain CN3911
 R;Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
 J. Bacteriol. 165, 21-27, 1986
 A;Title: Cloning, nucleotide sequencing, and expression of tetanus toxin
 fragment C in Escherichia coli.
 A;Reference number: A25194; MUID:86085672; PMID:3510187
 A;Accession: A25194
 A;Molecule type: DNA
 A;Residues: 743-1315 <FA2>
 A;Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921
 A;Accession: B25194
 A;Molecule type: protein
 A;Residues: 865-894 <FA3>
 R;Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
 Infect. Immun. 57, 3588-3593, 1989
 A;Title: Isolation, purification, and characterization of fragment B, the NH-2-
 terminal half of the heavy chain of tetanus toxin.
 A;Reference number: A60759; MUID:90035436; PMID:2478476
 A;Accession: A60759
 A;Molecule type: protein
 A;Residues: 461-475 <MAT>
 R;Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin,
 G.
 J. Immunol. 142, 394-402, 1989
 A;Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
 A;Reference number: JS0098; MUID:89093918; PMID:2463305
 A;Contents: annotation; epitope region
 R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.;
 DasGupta, B.R.; Montecucco, C.
 Nature 359, 832-835, 1992
 A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by
 proteolytic cleavage of synaptobrevin.
 A;Reference number: S27125; MUID:93063293; PMID:1331807
 A;Contents: annotation
 R;de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
 Eur. J. Biochem. 229, 61-69, 1995
 A;Title: Structural studies on the zinc-endopeptidase light chain of tetanus
 neurotoxin.
 A;Reference number: S69348; MUID:95262688; PMID:7744050
 A;Accession: S69348
 A;Molecule type: protein
 A;Residues: 2-31 <DEF>
 C;Comment: The source of this protein was an extrachromosomal plasmid.
 C;Comment: The precursor is cleaved by endogenous proteinase activity to form
 light (fragment A) and heavy (fragment B.C) chains that are covalently linked by
 an interchain disulfide bond (the individual chains are not toxic when
 separated). The amino end of the heavy chain (fragment B) can be separated from
 the carboxyl end (fragment C) by papain.
 C;Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to
 gangliosides and may target the toxin to the motor end plate. Fragment A is a
 zinc-dependent endopeptidase.

C;Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized, and moves by retrograde transport up the axon into the spinal cord where it can move between postsynaptic and presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of synaptobrevin (vesicle-associated membrane protein 2).

C;Function:

A;Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2

C;Superfamily: tetanus toxin

C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc

F;2-457/Product: tentoxylisin light chain (fragment A) #status predicted <TTL>

F;461-1315/Product: tentoxylisin heavy chain (fragment B.C) #status experimental <TTH>

F;461-864/Domain: channel forming (fragment B) #status predicted <TXB>

F;865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>

F;233,237/Binding site: zinc (His) #status predicted

F;234/Active site: Glu #status predicted

Query Match 100.0%; Score 112; DB 1; Length 1315;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | | |
Db 947 FNNFTVSFWLRVPKVSASHLE 967

RESULT 2

S33411

botulinum neurotoxin type F - Clostridium barati

C;Species: Clostridium barati

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999

C;Accession: S33411; S31860

R;Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T.

FEMS Microbiol. Lett. 108, 175-182, 1993

A;Title: Nucleotide sequence of the gene coding for Clostridium barati type F neurotoxin: comparison with other clostridial neurotoxins.

A;Reference number: S33411; MUID:93252228; PMID:8486245

A;Accession: S33411

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1268 <THO>

A;Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139

C;Superfamily: tetanus toxin

C;Keywords: neurotoxin

Query Match 55.4%; Score 62; DB 2; Length 1268;
Best Local Similarity 64.3%; Pred. No. 0.073;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
: | | : | | | | : | | |
Db 922 YQNFSVSFWVRIPK 935

RESULT 3

S48110

neurotoxin type F - Clostridium botulinum (fragment)

C;Species: Clostridium botulinum

C;Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999

C;Accession: S48110

R;Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific identification of neurotoxin types B, E, and F.

A;Reference number: S48103; MUID:94013372; PMID:8408542

A;Accession: S48110

A;Status: preliminary; translation not shown

A;Molecule type: DNA

A;Residues: 1-366 <CAM>

A;Cross-references: EMBL:X70821; NID:g407792; PIDN:CAA50152.1; PID:g407793

C;Superfamily: tetanus toxin

C;Keywords: neurotoxin

Query Match 54.5%; Score 61; DB 2; Length 366;
Best Local Similarity 57.1%; Pred. No. 0.028;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
: ||::|||:|:|
Db 297 YQNFSISFWVRIPK 310

RESULT 4

S48109

neurotoxin type F - Clostridium botulinum (fragment)

C;Species: Clostridium botulinum

C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999

C;Accession: S48109

R;Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific identification of neurotoxin types B, E, and F.

A;Reference number: S48103; MUID:94013372; PMID:8408542

A;Accession: S48109

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-369 <CAM>

A;Cross-references: EMBL:X70820; NID:g407790; PIDN:CAA50151.1; PID:g407791

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

C;Superfamily: tetanus toxin

Query Match 54.5%; Score 61; DB 2; Length 369;
Best Local Similarity 57.1%; Pred. No. 0.029;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
: ||::|||:|:|
Db 297 YQNFSISFWVRIPK 310

RESULT 5

I40813

neurotoxin type F - Clostridium botulinum

C;Species: Clostridium botulinum

C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999

C;Accession: I40813; S48108

R;East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, D.E.

FEMS Microbiol. Lett. 96, 225-230, 1992

A;Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.

A;Reference number: I40644

A;Accession: I40813

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1274 <RES>

A;Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867

R;Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific identification of neurotoxin types B, E, and F.

A;Reference number: S48103; MUID:94013372; PMID:8408542

A;Accession: S48108

A;Status: preliminary; translation not shown

A;Molecule type: DNA

A;Residues: 634-1002 <CAM>

A;Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789

C;Superfamily: tetanus toxin

C;Keywords: neurotoxin

Query Match 54.5%; Score 61; DB 2; Length 1274;

Best Local Similarity 57.1%; Pred. No. 0.11;

Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14

: ||::|||:|:|

Db 930 YQNFSISFWVRIPK 943

RESULT 6

S39791

neurotoxin - Clostridium botulinum

C;Species: Clostridium botulinum

C;Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 16-Jul-1999

C;Accession: S39791

R;Campbell, K.; Collins, M.D.; East, A.K.

Biochim. Biophys. Acta 1216, 487-491, 1993

A;Title: Nucleotide sequence of the gene coding for Clostridium botulinum (Clostridium argentinense) type G neurotoxin: genealogical comparison with other clostridial neurotoxins.

A;Reference number: S39791; MUID:94092745; PMID:8268233

A;Accession: S39791

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1297 <CAM>

A;Cross-references: EMBL:X74162; NID:g441275; PIDN:CAA52275.1; PID:g441276

C;Superfamily: tetanus toxin

C; Keywords: neurotoxin

Query Match 54.5%; Score 61; DB 2; Length 1297;
Best Local Similarity 38.1%; Pred. No. 0.11;
Matches 8; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 FNNFTVSFWLRVPKVSASHLE 21
      |:|:|::|:| | | : : ::
Db     930 FDNFSINFWVRTPKYNNNDIO 950
```

RESULT 7

BTCLAB

bontoxilysin (EC 3.4.24.69) A precursor - Clostridium botulinum

N;Alternate names: botulinum neurotoxin type A

C;Species: Clostridium botulinum

C;Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text change 18-Jun-1999

C;Accession: A35294; S09492; S68220; A33401; A53884; A60025; A27000

R; Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990

A;Title: The complete sequence of botulinum neurotoxin type A and comparison with other clostridial neurotoxins.

A;Reference number: A35294; MUID:90264400; PMID:2160960

A;Accession: A35294

A;Molecule type: DNA

A;Residues: 1-1296 <BIN>

A;Cross-references: GB:M30196; NID:g144864; PIDN:AAA23262.1; PID:g144865

A; Experimental source: strain 62A, subtype A

R;Thompson, D.E.; Brehm, J.K.; Oultram, J.D.; Swinfield, T.J.; Shone, C.C.;
Atkinson, T.; Melling, J.; Minton, N.P.

Eur. J. Biochem. 189, 73-81, 1990

A;Title: The complete amino acid sequence of the Clostridium botulinum type A neurotoxin, deduced by nucleotide sequence analysis of the encoding gene.

A;Reference number: S09492; MUID:90235864; PMID:2185020

A;Accession: S09492

A;Molecule type: DNA

A;Residues: 1, 'Q', 3-26, 'V', 28-1296 <THO>

A; Cross-references: EMBL:X52066; NID:g40381; PIDN:CAA36289.1; PID:g40382

A; Experimental source: NCTC 2916

R; Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.
FEBS Lett. 376, 41-44, 1995

A;Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin components of *Clostridium botulinum* type A progenitor toxins.

A;Reference number: S67988; MUID:96096783; PMID:8521962

A;Accession: 568220

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-12 <FUJ>

A;Cross-references: EMBL:D67030; DDBJ:D50421; NID:q2160224

R;Betley, M.J.; Somers, E.; DasGupta, B.R.

Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989

A;Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-terminal encoding region.

A;Reference number: A33401; MUID:89350959; PMID:2669749

A;Accession: A33401

A;Molecule type: DNA

A;Residues: 1-35 <BET>

A;Cross-references: GB:M27892; NID:g144880; PIDN:AAA23269.1; PID:g551776
 R;Gimenez, J.A.; DasGupta, B.R.
 J. Protein Chem. 12, 351-363, 1993
 A;Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42, and 18 kD fragments.
 A;Reference number: A53884; MUID:94000342; PMID:8397793
 A;Accession: A53884
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 867-880;1148-1217,'Y',1219 <GIM>
 A;Experimental source: strain Hall
 A;Note: sequence extracted from NCBI backbone (NCBIP:139159); sequence modified after extraction from NCBI backbone
 R;DasGupta, B.R.; Dekleva, M.L.
 Biochimie 72, 661-664, 1990
 A;Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and around the nicking site.
 A;Reference number: A60025; MUID:91120847; PMID:2126206
 A;Accession: A60025
 A;Molecule type: protein
 A;Residues: 2-6;445-453,'X',455-457 <DAS1>
 R;DasGupta, B.R.; Foley, J.; Niece, R.
 Biochemistry 26, 4162, 1987
 A;Title: Partial sequence of the light chain of botulinum neurotoxin type A.
 A;Reference number: A27000
 A;Accession: A27000
 A;Molecule type: protein
 A;Residues: 2-47 <DAS2>
 R;Binz, T.; Blasi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.; Niemann, H.
 J. Biol. Chem. 269, 1617-1620, 1994
 A;Title: Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.
 A;Reference number: A49708; MUID:94124495; PMID:8294407
 A;Contents: annotation
 C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synapses. This toxin is activated by cleavage into two chains linked by a disulfide bond.
 C;Genetics:
 A;Gene: atx; botA
 C;Function:
 A;Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associated 25K protein (SNAP-25)
 C;Superfamily: tetanus toxin
 C;Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
 F;2-444/Product: bontoxilysin A light chain #status experimental <LGHT>
 F;445-1296/Product: bontoxilysin A heavy chain #status experimental <HVY>
 F;223,227/Binding site: zinc (His) #status predicted
 F;224/Active site: Glu #status predicted

Query Match 52.7%; Score 59; DB 1; Length 1296;
 Best Local Similarity 57.1%; Pred. No. 0.24;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVVK 14
 : ||: |||:|:|
 Db 938 YENFSTSFWIRIPK 951

RESULT 8

A48940

bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum

N;Alternate names: botulinum neurotoxin type B (BoNT/B)

C;Species: Clostridium botulinum

C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999

C;Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574

R;Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.

Appl. Environ. Microbiol. 58, 2345-2354, 1992

A;Title: Molecular cloning of the Clostridium botulinum structural gene encoding the type B neurotoxin and determination of its entire nucleotide sequence.

A;Reference number: A48940; MUID:92384550; PMID:1514783

A;Accession: A48940

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1291 <WHE>

A;Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735

A;Experimental source: type B, Danish

A;Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBIP:112081); this publication is not cited in GenBank entry CLOBOTB, release 103

R;Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific identification of neurotoxin types B, E, and F.

A;Reference number: S48103; MUID:94013372; PMID:8408542

A;Accession: S48105

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 634-994 <CAM>

A;Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783

A;Experimental source: proteolytic type B, strain NCTC 7273

R;Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.

submitted to the EMBL Data Library, April 1992

A;Description: Partial amino acid sequence of botulinum neurotoxin type B and comparison to other Clostridial neurotoxins.

A;Reference number: S21575

A;Accession: S21575

A;Molecule type: DNA

A;Residues: 36-217, 'G', 219-224, 'S', 226-246 <SZA>

A;Cross-references: EMBL:Z11934; NID:g40383; PIDN:CAA77991.1; PID:g40384

R;Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars, K.; Poulain, B.; Tauc, L.; Niemann, H.

J. Biol. Chem. 267, 14721-14729, 1992

A;Title: Minimal essential domains specifying toxicity of the light chains of tetanus toxin and botulinum neurotoxin type A.

A;Reference number: A42871; MUID:92340509; PMID:1634516

A;Accession: A42871

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-313, 'S', 315-451 <KUR>

A;Experimental source: strain Okra

A;Note: sequence extracted from NCBI backbone (NCBIP:109365)

R;DasGupta, B.R.; Datta, A.

Biochimie 70, 811-817, 1988

A;Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with tetanus toxin.

A;Reference number: S07155; MUID:89000987; PMID:3139097

A;Accession: S07155

A;Molecule type: protein

A;Residues: 2-29,'M',31-45 <DAS>

A;Accession: S08562

A;Molecule type: protein

A;Residues: 442-463,'R',465-467 <DA2>

R;Schmidt, J.J.; Sathyamoorthy, V.; DasGupta, B.R.

Arch. Biochem. Biophys. 238, 544-548, 1985

A;Title: Partial amino acid sequences of botulinum neurotoxins types B and E.

A;Reference number: S07128; MUID:85197963; PMID:3888113

A;Accession: S07128

A;Status: preliminary

A;Molecule type: protein

A;Residues: 2-16 <SCH1>

A;Accession: S08573

A;Status: preliminary

A;Molecule type: protein

A;Residues: 2-17 <SCH2>

A;Accession: S08574

A;Status: preliminary

A;Molecule type: protein

A;Residues: 442-459 <SCH3>

R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.;

DasGupta, B.R.; Montecucco, C.

Nature 359, 832-835, 1992

A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin.

A;Reference number: S27125; MUID:93063293; PMID:1331807

A;Contents: annotation

C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synapses. This toxin is activated by cleavage into two chains linked by a disulfide bond.

C;Genetics:

A;Gene: bont/b

C;Function:

A;Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2

C;Superfamily: tetanus toxin

C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc

F;2-441/Product: bontoxilysin B light chain #status experimental <LGHT>

F;442-1291/Product: bontoxilysin B heavy chain #status experimental <HVY>

F;230,234/Binding site: zinc (His) #status predicted

F;231/Active site: Glu #status predicted

Query Match 51.8%; Score 58; DB 1; Length 1291;

Best Local Similarity 64.3%; Pred. No. 0.34;

Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14

| :|:|||||:|:|

Db 923 FLDFSVSFWIRIPK 936

RESULT 9

I40631

non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum

C;Species: Clostridium botulinum

C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999

C;Accession: I40631; S48103; S48104; S36015

R;Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.

Curr. Microbiol. 28, 101-110, 1994

A;Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulinum type B neurotoxin: comparison with other clostridial neurotoxins.

A;Reference number: I40631; MUID:94122659; PMID:7764370

A;Accession: I40631

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1291 <RES>

A;Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149

R;Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific identification of neurotoxin types B, E, and F.

A;Reference number: S48103; MUID:94013372; PMID:8408542

A;Accession: S48103

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 634-761,'E',763-841,'M',843,'T',845,'N',847-994 <CAM1>

A;Cross-references: EMBL:X70814; NID:g407778; PIDN:CAA50145.1; PID:g407779

A;Experimental source: non-proteolytic strain 2129B (Scott)

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

A;Accession: S48104

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 634-843,'T',845,'N',847-994 <CAM2>

A;Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781

A;Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)

C;Comment: Botulinum neurotoxin type B in these strains may possess a capable catalytic site but are nonetheless inactive.

C;Genetics:

A;Gene: bont/b

C;Superfamily: tetanus toxin

C;Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc

F;2-441/Product: botulinum neurotoxin type B light chain #status predicted <LGHT>

F;442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVY>

F;230,234/Binding site: zinc (His) #status predicted

F;231/Active site: Glu #status predicted

Query Match 51.8%; Score 58; DB 2; Length 1291;

Best Local Similarity 64.3%; Pred. No. 0.34;

Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVVK 14

| :|:|||||:|:|

Db 923 FLDFSVSFWIRIPK 936

RESULT 10

S48106

neurotoxin type E - Clostridium botulinum (fragment)

C;Species: Clostridium botulinum

C;Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999

C;Accession: S48106

R;Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A;Title: Gene probes for identification of the botulinum neurotoxin gene and specific identification of neurotoxin types B, E, and F.

A;Reference number: S48103; MUID:94013372; PMID:8408542

A;Accession: S48106

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-367 <CAM>

A;Cross-references: EMBL:X70818; NID:g407784; PIDN:CAA50149.1; PID:g407785

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

C;Superfamily: tetanus toxin

C;Keywords: neurotoxin

Query Match 50.0%; Score 56; DB 2; Length 367;
Best Local Similarity 53.8%; Pred. No. 0.19;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVP 13
: ||::|||:|:
Db 297 YKNFSISFWVRIP 309

RESULT 11

JH0256

botulinum neurotoxin type E precursor - Clostridium butyricum

C;Species: Clostridium butyricum

C;Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999

C;Accession: JH0256; S16145

R;Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.

Biochem. Biophys. Res. Commun. 183, 107-113, 1992

A;Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum type E (strain Beluga) and Clostridium butyricum (strains ATCC 43181 and ATCC 43755).

A;Reference number: JH0256; MUID:92181428; PMID:1543481

A;Accession: JH0256

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-27,'E',29-1251 <POU>

A;Cross-references: EMBL:X62088; NID:g40379

A;Experimental source: strains ATCC 43181 and ATCC 43755

R;Fujii, N.; Kimura, K.; Yashiki, T.; Indoh, T.; Murakami, T.; Tsuzuki, K.; Yokosawa, N.; Oguma, K.

J. Gen. Microbiol. 137, 519-525, 1991

A;Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E toxin gene from Clostridium butyricum strain BL6340.

A;Reference number: S16145; MUID:91237316; PMID:2033376

A;Accession: S16145

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-229,'M',231-252 <FUJ>
A;Cross-references: EMBL:X53180; NID:g40407; PIDN:CAA37321.1; PID:g40408
A;Experimental source: strain BL6340
C;Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter release at synaptic junctions.
C;Comment: The heavy chain mediates the binding of toxin to cell receptors while the light chain appears to enter target cells.
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin
F;2-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>
F;423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEA>
F;412-426/Disulfide bonds: #status predicted

Query Match 50.0%; Score 56; DB 2; Length 1251;
Best Local Similarity 53.8%; Pred. No. 0.72;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVP 13
: ||::|||:|:
Db 912 YKNFSISFWVRIP 924

RESULT 12

S21178

botulinum neurotoxin type E precursor - Clostridium botulinum

C;Species: Clostridium botulinum

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Oct-1999

C;Accession: S21178; S48107; JH0257; B35294; A60027; S18111

R;Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Atkinson, T.; Minton, N.P.

Eur. J. Biochem. 204, 657-667, 1992

A;Title: The complete amino acid sequence of the Clostridium botulinum type-E neurotoxin, derived by nucleotide-sequence analysis of the encoding gene.

A;Reference number: S21178; MUID:92174922; PMID:1541280

A;Accession: S21178

A;Molecule type: DNA

A;Residues: 1-1252 <WHE>

A;Cross-references: EMBL:X62683; NID:g40397; PIDN:CAA44558.1; PID:g40398

R;Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific identification of neurotoxin types B, E, and F.

A;Reference number: S48103; MUID:94013372; PMID:8408542

A;Accession: S48107

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 616-982 <CAM>

A;Cross-references: EMBL:X70815; NID:g407786; PIDN:CAA50146.1; PID:g407787

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

R;Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.

Biochem. Biophys. Res. Commun. 183, 107-113, 1992

A;Title: Sequences of the botulinal neurotoxin E derived from Clostridium botulinum type E (strain Beluga) and Clostridium butyricum (strains ATCC 43181 and ATCC 43755).

A;Reference number: JH0256; MUID:92181428; PMID:1543481

A;Accession: JH0257

A;Status: nucleic acid sequence not shown
 A;Molecule type: DNA
 A;Residues: 1-176,'R',178-197,'C',199-339,'R',341-772,'I',774-962,'FE',965-966,'R',968-1195,1197-1252 <POU>
 A;Cross-references: EMBL:X62089; NID:g40393; PIDN:CAA43999.1; PID:g40394
 A;Experimental source: strain Beluga
 R;Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
 J. Biol. Chem. 265, 9153-9158, 1990
 A;Title: The complete sequence of botulinum neurotoxin type A and comparison with other clostridial neurotoxins.
 A;Reference number: A35294; MUID:90264400; PMID:2160960
 A;Accession: B35294
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-176,'R',178-252 <BIN>
 A;Experimental source: strain Beluga
 R;Gimenez, J.A.; DasGupta, B.R.
 Biochimie 72, 213-217, 1990
 A;Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals the site trypsin nicks and homology with tetanus neurotoxin.
 A;Reference number: A60027; MUID:90344918; PMID:2116911
 A;Accession: A60027
 A;Molecule type: protein
 A;Residues: 420-427 <GIM>
 A;Experimental source: strain Beluga
 A;Note: this fragment was generated by proteolysis with Lys-C rather than with trypsin
 C;Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit neurotransmitter release at various synapses.
 C;Comment: The heavy chain mediates the binding of toxin to cell receptors while the light chain appears to enter target cells.
 C;Superfamily: tetanus toxin
 C;Keywords: neurotoxin
 F;1-422/Product: botulinum neurotoxin type E light chain #status predicted <LCH>
 F;423-1252/Product: botulinum neurotoxin type E heavy chain #status predicted <HCH>
 F;412-426/Disulfide bonds: #status predicted

Query Match 50.0%; Score 56; DB 2; Length 1252;
 Best Local Similarity 53.8%; Pred. No. 0.72;
 Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVP 13
 : ||::|||:|:
 Db 912 YKNFSISFWVRIP 924

RESULT 13
 I40645
 botulinum neurotoxin type A - Clostridium botulinum
 C;Species: Clostridium botulinum
 C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
 C;Accession: I40645
 R;Willems, A.; East, A.K.; Lawson, P.A.; Collins, M.D.
 Res. Microbiol. 144, 547-556, 1993

A;Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A associated with infant botulism: comparison with other clostridial neurotoxins.

A;Reference number: I40645; MUID:94143603; PMID:8310180

A;Accession: I40645

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1296 <RES>

A;Cross-references: EMBL:X73423; NID:g507070; PIDN:CAA51824.1; PID:g507071

C;Superfamily: tetanus toxin

C;Keywords: neurotoxin

Query Match 50.0%; Score 56; DB 2; Length 1296;
Best Local Similarity 50.0%; Pred. No. 0.74;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
: ||: |||::||
Db 938 YENFSTSFWIKIPK 951

RESULT 14

S23158

nucleocapsid protein - Impatiens necrotic spot virus

C;Species: Impatiens necrotic spot virus

C;Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 20-Sep-1999

C;Accession: S23158

R;de Haan, P.; de Avila, A.C.; Kormelink, R.; Westerbroek, A.; Gielen, J.J.L.; Peters, D.; Goldbach, R.

FEBS Lett. 306, 27-32, 1992

A;Title: The nucleotide sequence of the S RNA of Impatiens necrotic spot virus, a novel tospovirus.

A;Reference number: S23158; MUID:92331780; PMID:1385787

A;Accession: S23158

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1-449 <HAA>

A;Cross-references: EMBL:X66972; NID:g60488; PIDN:CAA47382.1; PID:g60489

A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 1 in having 104-Val, and 318-Met

C;Superfamily: tomato spotted wilt virus nonstructural protein NS

C;Keywords: nucleocapsid

Query Match 46.4%; Score 52; DB 2; Length 449;
Best Local Similarity 45.0%; Pred. No. 1.1;
Matches 9; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

Qy 2 NN--FTVSFWLRVPKVSASH 19
|| | :| |:|:|:| :| :
Db 240 NNKPFKISLWMRIPKIMKSN 259

RESULT 15

MNVUWC

nonstructural protein NS - tomato spotted wilt virus (strain CPNH1)

C;Species: tomato spotted wilt virus

C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jun-2000

C;Accession: JQ0547
R;de Haan, P.; Wagemakers, L.; Peters, D.; Goldbach, R.
J. Gen. Virol. 71, 1001-1007, 1990
A;Title: The S RNA segment of tomato spotted wilt virus has an ambisense character.
A;Reference number: JQ0547; MUID:90264829; PMID:1693160
A;Accession: JQ0547
A;Molecule type: genomic RNA
A;Residues: 1-464 <DEH>
A;Cross-references: DDBJ:D00645; NID:g222685; PIDN:BAA00540.1; PID:g222686
C;Genetics:
A;Gene: NS
A;Map position: segment S
C;Superfamily: tomato spotted wilt virus nonstructural protein NS
C;Keywords: glycoprotein; nonstructural protein
F;132,210,270,291,381/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 46.4%; Score 52; DB 1; Length 464;
Best Local Similarity 64.3%; Pred. No. 1.1;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 NNFTVSFWLRVPKV 15
::| :| |||||
Db 239 SHFKLSLWLRVPKV 252

Search completed: December 13, 2003, 13:30:54
Job time : 2.30769 secs

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 0.717949 Seconds
(without alignments)
1375.531 Million cell updates/sec

Title: US-09-785-215-6
Perfect score: 112
Sequence: 1 FNNFTVSFWLRVPKVSASHLE 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	112	100.0	1314	1	TETX_CLOTE	P04958 clostridium
2	61	54.5	1274	1	BXF_CLOBO	P30996 clostridium
3	61	54.5	1296	1	BXG_CLOBO	Q60393 clostridium
4	59	52.7	1295	1	BXA1_CLOBO	P10845 clostridium
5	58	51.8	1290	1	BXB_CLOBO	P10844 clostridium
6	57	50.9	1051	1	VP2_AHSV6	O71024 african hor
7	56	50.0	1250	1	BXE_CLOBO	Q00496 clostridium
8	56	50.0	1250	1	BXE_CLOBU	P30995 clostridium
9	56	50.0	1295	1	BXA2_CLOBO	Q45894 clostridium
10	52	46.4	449	1	VNSS_INSVN	Q01811 impatiens n
11	52	46.4	464	1	VNSS_TSWV1	P26002 tomato spot
12	52	46.4	467	1	VNSS_TSWVL	P26003 tomato spot
13	51	45.5	1196	1	BXCN_CLOBO	P46081 clostridium
14	48	42.9	1290	1	BXC1_CLOBO	P18640 clostridium
15	47.5	42.4	1276	1	BXD_CLOBO	P19321 clostridium
16	47	42.0	760	1	AMY_CLOAB	P23671 clostridium
17	46	41.1	1162	1	BXEN_CLOBO	P46082 clostridium

18	46	41.1	1162	1	BXEN_CLOBU	Q06366	clostridium
19	45	40.2	639	1	CGMA_RHIME	P72302	rhizobium m
20	44	39.3	241	1	CRTA_RHOCA	P17055	rhodobacter
21	43	38.4	152	1	YOBD_ECOLI	P76263	escherichia
22	43	38.4	867	1	RRPO_BYDV1	P29044	barley yell
23	42	37.5	208	1	FTSQ_STRGR	P45503	streptomyce
24	42	37.5	789	1	AI2M_YEAST	P03876	saccharomyc
25	42	37.5	1337	1	YDM5_SCHPO	P87136	schizosacch
26	42	37.5	1539	1	SMCY_HUMAN	Q9by66	homo sapien
27	42	37.5	1548	1	SMCY_MOUSE	Q62240	mus musculu
28	42	37.5	1554	1	SMCX_MOUSE	P41230	mus musculu
29	42	37.5	1560	1	SMCX_HUMAN	P41229	homo sapien
30	41.5	37.1	237	1	YBS0_YEAST	P38242	saccharomyc
31	41.5	37.1	728	1	YJ89_YEAST	P47156	saccharomyc
32	41	36.6	352	1	CITC_ECOLI	P77390	escherichia
33	40.5	36.2	230	1	FIMC_SALTY	P37923	salmonella
34	40.5	36.2	535	1	YA98_SCHPO	Q09786	schizosacch
35	40	35.7	116	1	NU3M_PAROL	P92817	paralichthy
36	40	35.7	118	1	YMX2_YEAST	Q04276	saccharomyc
37	40	35.7	269	1	PARA_MYCGE	P47706	mycoplasma
38	40	35.7	288	1	PEF1_SCHPO	O74456	schizosacch
39	40	35.7	385	1	G22C_DROME	P58952	drosophila
40	40	35.7	390	1	DACD_SALTY	P37604	salmonella
41	40	35.7	396	1	YJE8_YEAST	P47049	saccharomyc
42	40	35.7	402	1	LE22_METKA	Q8tw29	methanopyru
43	40	35.7	453	1	NRAM_IAWIL	P03470	influenza a
44	40	35.7	454	1	NRAM_IAPUE	P03468	influenza a
45	40	35.7	469	1	NRAM_IAPAR	P06819	influenza a

ALIGNMENTS

RESULT 1

TETX_CLOTE

ID TETX_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) [Contains:
DE Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy
DE chain (Tetanus toxin chain H)].
GN TETX OR CTP60.
OS Clostridium tetani.
OG Plasmid pE88, and Plasmid 75 Kbp.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudel M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins.";

RL EMBO J. 5:2495-2502(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CN3911; PLASMID=75 Kbp;
 RX MEDLINE=87040747; PubMed=3774547;
 RA Fairweather N.F., Lyness V.A.;
 RT "The complete nucleotide sequence of tetanus toxin.";
 RL Nucleic Acids Res. 14:7809-7812(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Massachusetts / E88; PLASMID=pE88;
 RX MEDLINE=22457253; PubMed=12552129;
 RA Brueggemann H., Baeumer S., Fricke W.F., Wiezer A., Liesegang H.,
 RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
 RA Gottschalk G.;
 RT "The genome sequence of Clostridium tetani, the causative agent of
 RT tetanus disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
 RN [4]
 RP SEQUENCE OF 742-1314 FROM N.A.
 RC PLASMID=75 Kbp;
 RX MEDLINE=86085672; PubMed=3510187;
 RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
 RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
 RT fragment C in Escherichia coli.";
 RL J. Bacteriol. 165:21-27(1986).
 RN [5]
 RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 RX MEDLINE=90201034; PubMed=2108021;
 RA Krieglstein K., Henschen A., Weller U., Habermann E.;
 RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
 RT in tetanus toxin.";
 RL Eur. J. Biochem. 188:39-45(1990).
 RN [6]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=92037649; PubMed=1935979;
 RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
 RT "Limited proteolysis of tetanus toxin. Relation to activity and
 RT identification of cleavage sites.";
 RL Eur. J. Biochem. 202:41-51(1991).
 RN [7]
 RP IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93010948; PubMed=1396558;
 RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
 RA Montecucco C.;
 RT "Tetanus toxin is a zinc protein and its inhibition of
 RT neurotransmitter release and protease activity depend on zinc.";
 RL EMBO J. 11:3577-3583(1992).
 RN [8]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
 RT by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 RN [9]

RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
 RX MEDLINE=97475217; PubMed=9334741;
 RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
 RA Sax M.;
 RT "Structure of the receptor binding fragment HC of tetanus
 RT neurotoxin.";
 RL Nat. Struct. Biol. 4:788-792(1997).
 CC -!- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
 CC BOND OF SYNAPTOBREVIN-2.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in
 CC synaptobrevin 2.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
 CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
 CC AND ARE NON-TOXIC AFTER SEPARATION.
 CC -!- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
 CC GANGLIOSIDE RECEPTORS.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X04436; CAA28033.1; -.
 DR EMBL; X06214; CAA29564.1; -.
 DR EMBL; AF528097; AA037454.1; -.
 DR EMBL; M12739; AAA23282.1; -.
 DR PIR; A25689; BTCLTN.
 DR PDB; 1AF9; 29-APR-98.
 DR PDB; 1A8D; 14-OCT-98.
 DR PDB; 1DOH; 27-MAR-00.
 DR PDB; 1DFQ; 24-MAR-00.
 DR PDB; 1DIW; 24-MAR-00.
 DR PDB; 1DLL; 24-MAR-00.
 DR PDB; 1FV3; 05-SEP-01.
 DR MEROPS; M27.001; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
 KW 3D-structure; Complete proteome.
 FT INIT_MET 0 0
 FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.
 FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.
 FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).

FT	ACT_SITE	233	233
FT	METAL	236	236
FT	TRANSMEM	226	246
FT	TRANSMEM	669	689
FT	DISULFID	438	466
FT	DISULFID	1076	1092
FT	HELIX	876	882
FT	TURN	883	883
FT	STRAND	884	891
FT	TURN	892	893
FT	STRAND	894	897
FT	STRAND	904	907
FT	TURN	909	910
FT	STRAND	912	915
FT	STRAND	920	925
FT	TURN	928	929
FT	STRAND	932	935
FT	HELIX	938	940
FT	TURN	941	946
FT	STRAND	949	956
FT	HELIX	962	968
FT	TURN	969	970
FT	STRAND	972	977
FT	STRAND	980	981
FT	HELIX	983	985
FT	STRAND	987	995
FT	TURN	996	997
FT	STRAND	998	1004
FT	TURN	1006	1007
FT	STRAND	1010	1016
FT	STRAND	1020	1020
FT	TURN	1021	1022
FT	STRAND	1031	1037
FT	TURN	1039	1040
FT	STRAND	1042	1047
FT	TURN	1048	1049
FT	STRAND	1050	1056
FT	TURN	1058	1059
FT	STRAND	1068	1074
FT	TURN	1079	1080
FT	STRAND	1082	1091
FT	HELIX	1097	1105
FT	TURN	1106	1107
FT	STRAND	1112	1112
FT	STRAND	1114	1114
FT	TURN	1116	1117
FT	STRAND	1120	1120
FT	STRAND	1122	1122
FT	TURN	1123	1124
FT	STRAND	1127	1131
FT	HELIX	1132	1134
FT	TURN	1135	1136
FT	STRAND	1137	1141
FT	TURN	1144	1145
FT	STRAND	1148	1152
FT	STRAND	1155	1158
FT	TURN	1159	1162

BY SIMILARITY.
 ZINC (CATALYTIC) (BY SIMILARITY).
 POTENTIAL.
 POTENTIAL.
 INTERCHAIN.

FT	STRAND	1163	1166
FT	STRAND	1173	1178
FT	TURN	1184	1185
FT	STRAND	1188	1188
FT	STRAND	1190	1190
FT	TURN	1191	1192
FT	STRAND	1193	1201

Query Match 100.0%; Score 112; DB 1; Length 1314;
 Best Local Similarity 100.0%; Pred. No. 2.4e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 FNNFTVSFWLRVPKVSASHLE 21
Db	946 FNNFTVSFWLRVPKVSASHLE 966

RESULT 2

BXF_CLOBO

ID BXF_CLOBO STANDARD; PRT; 1274 AA.
 AC P30996;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BoNT/F)
 DE (Bontoxilysin F).
 GN BOTF.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 23387;
 RX MEDLINE=93012902; PubMed=1398040;
 RA East A.K., Richardson P.T., Allaway D., Collins M.D.,
 RA Roberts T.A., Thompson D.E.;
 RT "Sequence of the gene encoding type F neurotoxin of Clostridium
 RT botulinum.";
 RL FEMS Microbiol. Lett. 75:225-230(1992).
 RN [2]
 RP SEQUENCE OF 1-64 FROM N.A.
 RC STRAIN=Hobbs FT10;
 RX MEDLINE=94297488; PubMed=7764998;
 RA East A.K., Collins M.D.;
 RT "Conserved structure of genes encoding components of botulinum
 RT neurotoxin complex M and the sequence of the gene coding for the
 RT nontoxic component in nonproteolytic Clostridium botulinum type F.";
 RL Curr. Microbiol. 29:69-77(1994).
 RN [3]
 RP SEQUENCE OF 634-1002 FROM N.A.
 RX MEDLINE=94013372; PubMed=8408542;
 RA Campbell K., East A.K., Collins M.D.;
 RT "Gene probes for identification of the botulinal neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F.";
 RL J. Clin. Microbiol. 31:2255-2262(1993).
 RN [4]

RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94230352; PubMed=8175689;
RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA Roques B., Fykse E.M., Suedhof T.C., Jahn R., Niemann H.;
RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT F botulinial neurotoxins and tetanus toxin.";
RL J. Biol. Chem. 269:12764-12772(1994).

CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-|-LYS-59
CC BOND OF SYNAPTOBREVINS-1 AND -2.

CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.

CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M92906; AAA23263.1; -.
DR EMBL; S73676; AAC60475.1; -.
DR EMBL; X70820; CAA50151.1; -.
DR EMBL; X70816; CAA50147.1; -.
DR PIR; I40813; I40813.
DR PIR; S48109; S48109.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR006025; Zn_MTpeptdse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.

KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.

FT CHAIN 1 436 BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.
FT CHAIN 437 1274 BOTULINUM NEUROTOXIN F, HEAVY-CHAIN.
FT METAL 227 227 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 228 228 BY SIMILARITY.
FT METAL 231 231 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 429 445 INTERCHAIN (PROBABLE).
SQ SEQUENCE 1274 AA; 146709 MW; 5B99756A7438B921 CRC64;

Query Match 54.5%; Score 61; DB 1; Length 1274;
 Best Local Similarity 57.1%; Pred. No. 0.058;
 Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14
 : ||::|||:|:|
 Db 930 YQNFSISFWVRIPK 943

RESULT 3

BXG_CLOBO

ID BXG_CLOBO STANDARD; PRT; 1296 AA.
 AC Q60393;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BoNT/G)
 DE (Bontoxilysin G).
 GN BOTG.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=113 / 30;
 RX MEDLINE=94092745; PubMed=8268233;
 RA Campbell K., Collins M.D., East A.K.;
 RT "Nucleotide sequence of the gene coding for Clostridium botulinum
 RT (Clostridium argentinense) type G neurotoxin: genealogical comparison
 RT with other clostridial neurotoxins."
 RL Biochim. Biophys. Acta 1216:487-491(1993).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 CC heavy chain (H). The light chain has the pharmacological activity,
 CC while the N- and C-terminal of the heavy chain mediate channel
 CC formation and toxin binding, respectively.
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

RC STRAIN=62A;
 RX MEDLINE=90264400; PubMed=2160960;
 RA Binz B., Kuarzono H., Wille M., Frevent J., Wernars K., Niemann H.;
 RT "The complete sequence of botulinum neurotoxin type A and comparison
 RT with other clostridial neurotoxins.";
 RL J. Biol. Chem. 265:9153-9158(1990).
 RN [3]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN=62A;
 RX MEDLINE=97016817; PubMed=8863443;
 RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
 RT "Organization and phylogenetic interrelationships of genes encoding
 RT components of the botulinum toxin complex in proteolytic Clostridium
 RT botulinum types A, B, and F: evidence of chimeric sequences in the
 RT gene encoding the nontoxic nonhemagglutinin component.";
 RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
 RN [4]
 RP SEQUENCE OF 1-34 FROM N.A.
 RC STRAIN=Hall;
 RX MEDLINE=89350959; PubMed=2669749;
 RA Betley M.J., Somers E., Dasgupta B.R.;
 RT "Characterization of botulinum type A neurotoxin gene: delineation of
 RT the N-terminal encoding region.";
 RL Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
 RN [5]
 RP SEQUENCE OF 1-18 FROM N.A.
 RC STRAIN=Type A NIH;
 RX MEDLINE=96096783; PubMed=8521962;
 RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
 RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin
 RT components of Clostridium botulinum type A progenitor toxins.";
 RL FEBS Lett. 376:41-44(1995).
 RN [6]
 RP SEQUENCE OF 1-16.
 RX MEDLINE=84178501; PubMed=6370252;
 RA Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;
 RT "Partial amino acid sequence of the heavy and light chains of
 RT botulinum neurotoxin type A.";
 RL Biochem. Biophys. Res. Commun. 119:900-904(1984).
 RN [7]
 RP SEQUENCE OF 1-46.
 RA Dasgupta B.R., Foley J., Niece R.;
 RT "Partial sequence of the light chain of botulinum neurotoxin type A.";
 RL Biochemistry 26:4162-4162(1987).
 RN [8]
 RP SEQUENCE OF 1-5 AND 444-456.
 RX MEDLINE=91120847; PubMed=2126206;
 RA Dasgupta B.R., Dekleva M.L.;
 RT "Botulinum neurotoxin type A: sequence of amino acids at the
 RT N-terminus and around the nicking site.";
 RL Biochimie 72:661-664(1990).
 RN [9]
 RP SEQUENCE OF 448-464 AND 872-895.
 RX MEDLINE=89024662; PubMed=3178218;
 RA Sathymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;
 RT "Botulinum neurotoxin type A: cleavage of the heavy chain into two
 RT halves and their partial sequences.";

RL Arch. Biochem. Biophys. 266:142-151(1988).
 RN [10]
 RP SEQUENCE OF 448-482.
 RX MEDLINE=85285016; PubMed=3896784;
 RA Shone C.C., Hambleton P., Melling J.;
 RT "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
 RT and purification of two tryptic fragments. Proteolytic action near
 RT the COOH-terminus of the heavy subunit destroys toxin-binding
 RT activity.";
 RL Eur. J. Biochem. 151:75-82(1985).
 RN [11]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94063091; PubMed=8243676;
 RA Schiavo G., Santtuci A., Dasgupta B.R., Mehta P.P., Jontes J.,
 RA Benfenati F., Wilson M.C., Montecucco C.;
 RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
 RT COOH-terminal peptide bonds.";
 RL FEBS Lett. 335:99-103(1993).
 RN [12]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94124495; PubMed=8294407;
 RA Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RA Jahn R., Niemann H.;
 RT "Proteolysis of SNAP-25 by types E and A¹ botulinal neurotoxins.";
 RL J. Biol. Chem. 269:1617-1620(1994).
 RN [13]
 RP MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.
 RX MEDLINE=21556941; PubMed=11700044;
 RA Rigoni M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.;
 RT "Site-directed mutagenesis identifies active-site residues of the
 RT light chain of botulinum neurotoxin type a.";
 RL Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
 RN [14]
 RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
 RX MEDLINE=98455071; PubMed=9783750;
 RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;
 RT "Crystal structure of botulinum neurotoxin type A and implications
 RT for toxicity.";
 RL Nat. Struct. Biol. 5:898-902(1998).
 CC -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
 CC binds with high affinity to peripheral neuronal presynaptic
 CC membrane, is then internalized by receptor-mediated endocytosis.
 CC The C-terminus of the heavy chain (H) is responsible for the
 CC adherence of the toxin to the cell surface while the N-terminus
 CC mediates transport of the light chain from the endocytic vesicle
 CC to the cytosol. After translocation, the light chain (L)
 CC hydrolyzes the 197-Gln-| -Arg-198 bond in SNAP-25, thereby blocking
 CC neurotransmitter release. Inhibition of acetylcholine release
 CC results in flaccid paralysis, with frequent heart or respiratory
 CC failure.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit.
 CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 CC heavy chain (H).
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- PHARMACEUTICAL: Available under the name BOTOX (Allergan) for
 CC the treatment of strabismus and blepharospasm associated with
 CC dystonia and cervical dystonia. Also used for the treatment of
 CC hemifacial spasm and a number of other neurological disorders
 CC characterized by abnormal muscle contraction.
 CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -!- DATABASE: NAME=BOTOX product information Web site;
 CC WWW="http://www.botox.com/index.jsp?hp&productinfo".
 CC -!- DATABASE: NAME=Protein Spotlight;
 CC NOTE=Issue 19 of February 2002;
 CC WWW="http://www.expasy.org/spotlight/articles/sptlt019.html".

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X52066; CAA36289.1; -.
 DR EMBL; M30196; AAA23262.1; -.
 DR EMBL; X92973; CAA63551.1; -.
 DR EMBL; D67030; BAA11051.1; -.
 DR EMBL; M27892; AAA23269.1; -.
 DR PIR; A35294; BTCLAB.
 DR PDB; 3BTA; 01-OCT-99.
 DR MEROPS; M27.002; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
 KW Pharmaceutical; 3D-structure.
 FT INIT_MET 0 0
 FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
 FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
 FT METAL 222 222 ZINC (CATALYTIC).
 FT ACT_SITE 223 223
 FT METAL 226 226 ZINC (CATALYTIC).
 FT METAL 261 261 ZINC (CATALYTIC).
 FT DISULFID 429 453 INTERCHAIN.
 FT DISULFID 1234 1279
 FT TRANSMEM 626 646 POTENTIAL.
 FT TRANSMEM 655 675 POTENTIAL.
 FT VARIANT 26 26 V -> A.
 FT MUTAGEN 261 261 E->A: DRASTIC DECREASE IN ENZYMATI
 FT ACTIVITY.
 FT MUTAGEN 265 265 F->A: DECREASE IN ENZYMATI
 FT MUTAGEN 365 365 Y->A: DECREASE IN ENZYMATI
 FT CONFLICT 1 1 P -> Q (IN REF. 1).
 FT CONFLICT 479 479 E -> P (IN REF. 9).
 FT CONFLICT 875 875 T -> L (IN REF. 8).
 FT CONFLICT 891 891 S -> K (IN REF. 8).

SQ SEQUENCE 1295 AA; 149322 MW; 858342F754862579 CRC64;
Query Match 52.7%; Score 59; DB 1; Length 1295;
Best Local Similarity 57.1%; Pred. No. 0.13;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14
: ||: |||:|:|
Db 937 YENFSTSFWIRIPK 950

RESULT 5

BXB_CLOBO

ID BXB_CLOBO STANDARD; PRT; 1290 AA.
AC P10844; P10843;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)
DE (Bontoxilysin B).
GN BOTB.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92384550; PubMed=1514783;
RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
RA Minton N.P.;
RT "Molecular cloning of the Clostridium botulinum structural gene
RT encoding the type B neurotoxin and determination of its entire
RT nucleotide sequence.";
RL Appl. Environ. Microbiol. 58:2345-2354(1992).
RN [2]
RP SEQUENCE OF 35-245 FROM N.A.
RC STRAIN=NCTC 7273;
RA Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 633-993 FROM N.A.
RC STRAIN=NCTC 7273;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinal neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP SEQUENCE OF 1-44 AND 441-466.
RC STRAIN=657;
RX MEDLINE=89000987; PubMed=3139097;
RA Dasgupta B.R., Datta A.;
RT "Botulinum neurotoxin type B (strain 657): partial sequence and
RT similarity with tetanus toxin.";
RL Biochimie 70:811-817(1988).
RN [5]
RP SEQUENCE OF 1-16 AND 441-458.

RC STRAIN=OKRA;
 RX MEDLINE=85197963; PubMed=3888113;
 RA Schmidt J.J., Sathyamoorthy V., Dasgupta B.R.;
 RT "Partial amino acid sequences of botulinum neurotoxins types B and
 RT E.";
 RL Arch. Biochem. Biophys. 238:544-548(1985).
 RN [6]
 RP IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93054694; PubMed=1429690;
 RA Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
 RT "Botulinum neurotoxins are zinc proteins.";
 RL J. Biol. Chem. 267:23479-23483(1992).
 RN [7]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
 RT by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-|-PHE-77 BOND OF
 CC SYNAPTOBREVIN-2.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M81186; AAA23211.1; -.
 DR EMBL; Z11934; CAA77991.1; -.
 DR EMBL; X70817; CAA50148.1; -.
 DR PIR; A48940; A48940.
 DR PDB; 1EPW; 01-NOV-00.
 DR PDB; 1F31; 01-NOV-00.
 DR PDB; 1F82; 16-AUG-00.
 DR PDB; 1F83; 16-AUG-00.
 DR PDB; 1FQH; 06-DEC-00.

DR PDB; 1G9A; 13-NOV-02.
 DR PDB; 1G9B; 13-NOV-02.
 DR PDB; 1G9C; 13-NOV-02.
 DR PDB; 1G9D; 13-NOV-02.
 DR PDB; 1I1E; 21-NOV-01.
 DR MEROPS; M27.002; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
 KW 3D-structure.
 FT INIT_MET 0 0
 FT CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
 FT CHAIN 441 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
 FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 230 230 BY SIMILARITY.
 FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 436 445 INTERCHAIN (PROBABLE).
 FT CONFLICT 29 29 T -> M (IN REF. 4).
 FT CONFLICT 217 217 R -> G (IN REF. 2).
 FT CONFLICT 224 224 A -> S (IN REF. 2).
 FT CONFLICT 463 463 S -> R (IN REF. 4).
 SQ SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;

Query Match 51.8%; Score 58; DB 1; Length 1290;
 Best Local Similarity 64.3%; Pred. No. 0.18;
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
 | :|:||||:|:|
 Db 922 FLDFSVSFWIRIPK 935

RESULT 6

VP2_AHSV6

ID VP2_AHSV6 STANDARD; PRT; 1051 AA.
 AC 071024;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Outer capsid protein VP2.
 GN S2 OR L2.
 OS African horse sickness virus 6 (AHSV-6) (African horse sickness virus
 OS (serotype 6)).
 OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
 OX NCBI_TaxID=86060;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98278331; PubMed=9617769;
 RA Williams C.F., Inoue T., Lucus A.-M., Zanotto P., Roy P.;
 RT "The complete sequence of four major structural proteins of African
 RT horse sickness virus serotype 6: evolutionary relationships within
 RT and between the orbiviruses."
 RL Virus Res. 53:53-73(1998).

CC -!- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP5)
 CC WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE
 CC MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.
 CC -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF021235; AAC40994.1; -.
 DR InterPro; IPR001742; Orbi_VP2.
 DR Pfam; PF00898; Orbi_VP2; 1.
 KW Coat protein.
 SQ SEQUENCE 1051 AA; 122326 MW; 2B04DB9E389F4B5F CRC64;

Query Match 50.9%; Score 57; DB 1; Length 1051;
 Best Local Similarity 47.6%; Pred. No. 0.21;
 Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
 |: ||:| || |:| |||
 Db 636 FSKRFVSYWYRVEKITTKHLE 656

RESULT 7

BXE_CLOBO
 ID BXE_CLOBO STANDARD; PRT; 1250 AA.
 AC Q00496;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BoNT/E)
 DE (Bontoxilysin E).
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Beluga;
 RX MEDLINE=92181428; PubMed=1543481;
 RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
 RT "Sequences of the botulinum neurotoxin E derived from Clostridium
 RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
 RT ATCC 43181 and ATCC 43755).";
 RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92174922; PubMed=1541280;
 RA Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
 RT "The complete amino acid sequence of the Clostridium botulinum type-E
 RT neurotoxin, derived by nucleotide-sequence analysis of the encoding
 RT gene.";

RL Eur. J. Biochem. 204:657-667(1992).
 RN [3]
 RP SEQUENCE OF 1-251 FROM N.A.
 RX MEDLINE=90264400; PubMed=2160960;
 RA Binz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;
 RT "The complete sequence of botulinum neurotoxin type A and comparison
 RT with other clostridial neurotoxins.";
 RL J. Biol. Chem. 265:9153-9158(1990).
 RN [4]
 RP SEQUENCE OF 1-13.
 RX MEDLINE=85197963; PubMed=3888113;
 RA Schmidt J.J., Sathyamoorthy V., Dasgupta B.R.;
 RT "Partial amino acid sequences of botulinum neurotoxins types B and
 RT E.";
 RL Arch. Biochem. Biophys. 238:544-548(1985).
 RN [5]
 RP SEQUENCE OF 419-426.
 RX MEDLINE=90344918; PubMed=2116911;
 RA Gimenez J.A., Dasgupta B.R.;
 RT "Botulinum neurotoxin type E fragmented with endoproteinase Lys-C
 RT reveals the site trypsin nicks and homology with tetanus
 RT neurotoxin.";
 RL Biochimie 72:213-217(1990).
 RN [6]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94063091; PubMed=8243676;
 RA Schiavo G., Santtuci A., Dasgupta B.R., Mehta P.P., Jontes J.,
 RA Benfenati F., Wilson M.C., Montecucco C.;
 RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
 RT COOH-terminal peptide bonds.";
 RL FEBS Lett. 335:99-103(1993).
 RN [7]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94124495; PubMed=8294407;
 RA Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RA Jahn R., Niemann H.;
 RT "Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.";
 RL J. Biol. Chem. 269:1617-1620(1994).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-|-ILE-
 CC 181 BOND IN SNAP-25.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X62089; CAA43999.1; -.
 DR EMBL; X62683; CAA44558.1; -.
 DR PIR; S21178; S21178.
 DR HSSP; P10845; 3BTA.
 DR MEROPS; M27.002; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT INIT_MET 0 0
 FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
 FT CHAIN 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
 FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 411 425 INTERCHAIN (PROBABLE).
 FT CONFLICT 176 176 R -> G (IN REF. 2).
 FT CONFLICT 197 197 C -> S (IN REF. 2 AND 3).
 FT CONFLICT 339 339 R -> A (IN REF. 2).
 FT CONFLICT 772 772 I -> L (IN REF. 2).
 FT CONFLICT 962 963 FE -> LQ (IN REF. 2).
 FT CONFLICT 966 966 R -> A (IN REF. 2).
 FT CONFLICT 1194 1194 N -> NN (IN REF. 2).
 SQ SEQUENCE 1250 AA; 143712 MW; D9FCE26DDA041EB4 CRC64;

Query Match 50.0%; Score 56; DB 1; Length 1250;
 Best Local Similarity 53.8%; Pred. No. 0.38;
 Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVP 13
 : ||::|||:|:
 Db 911 YKNFSISFWVRIP 923

RESULT 8

BXE_CLOBU

ID BXE_CLOBU STANDARD; PRT; 1250 AA.
 AC P30995;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BoNT/E)
 DE (Bontoxilysin E).
 OS Clostridium butyricum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.
 OX NCBI_TaxID=1492;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 43181, and ATCC 43755;
 RX MEDLINE=92181428; PubMed=1543481;
 RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
 RT "Sequences of the botulinum neurotoxin E derived from Clostridium
 RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
 RT ATCC 43181 and ATCC 43755).";
 RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
 RN [2]
 RP SEQUENCE OF 1-251 FROM N.A.
 RC STRAIN=BL6340;
 RX MEDLINE=91237316; PubMed=2033376;
 RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
 RA Yokosawa N., Yashiki T., Oguma K.;
 RT "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
 RT type E toxin gene from Clostridium butyricum strain BL6340.";
 RL J. Gen. Microbiol. 137:519-525(1991).
 RN [3]
 RP SEQUENCE OF 1-48.
 RC STRAIN=5262;
 RA Gimenez J., Foley J., Dasgupta B.R.;
 RT "Neurotoxin type E from Clostridium botulinum and C. butyricum;
 RT partial sequence and comparison.";
 RL FASEB J. 2:A1750-A1750(1988).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates..
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X62088; CAA43998.1; -.
 DR EMBL; X53180; CAA37321.1; -.
 DR PIR; JH0256; JH0256.

DR HSSP; P10845; 3BTA.
 DR MEROPS; M27.002; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT INIT_MET 0 0
 FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
 FT CHAIN 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
 FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 411 425 INTERCHAIN (PROBABLE).
 FT CONFLICT 229 229 K -> M (IN REF. 2).
 SQ SEQUENCE 1250 AA; 143265 MW; 8171B5B2C2312857 CRC64;

Query Match 50.0%; Score 56; DB 1; Length 1250;
 Best Local Similarity 53.8%; Pred. No. 0.38;
 Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVP 13
 : ||::|||:|:
 Db 911 YKNFSISFWVRIP 923

RESULT 9

BXA2_CLOBO

ID BXA2_CLOBO STANDARD; PRT; 1295 AA.
 AC Q45894; P77780;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
 DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
 DE chain; Botulinum neurotoxin A, heavy-chain].
 GN BOTA OR BNA OR ATX.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Kyoto-F;
 RX MEDLINE=94143603; PubMed=8310180;
 RA Willems A., East A.K., Lawson P.A., Collins M.D.;
 RT "Sequence of the gene coding for the neurotoxin of Clostridium
 RT botulinum type A associated with infant botulism: comparison with
 RT other clostridial neurotoxins."
 RL Res. Microbiol. 144:547-556(1993).
 RN [2]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN=Kyoto-F;
 RX MEDLINE=97016817; PubMed=8863443;
 RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;

RT "Organization and phylogenetic interrelationships of genes encoding
RT components of the botulinum toxin complex in proteolytic Clostridium
RT botulinum types A, B, and F: evidence of chimeric sequences in the
RT gene encoding the nontoxic nonhemagglutinin component.";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).

CC -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC binds with high affinity to peripheral neuronal presynaptic
CC membrane, is then internalized by receptor-mediated endocytosis.
CC The C-terminus of the heavy chain (H) is responsible for the
CC adherence of the toxin to the cell surface while the N-terminus
CC mediates transport of the light chain from the endocytic vesicle
CC to the cytosol. After translocation, the light chain (L)
CC hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking
CC neurotransmitter release. Inhibition of acetylcholine release
CC results in flaccid paralysis, with frequent heart or respiratory
CC failure (By similarity).

CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.

CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC heavy chain (H) (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X73423; CAA51824.1; -.
DR EMBL; X87974; CAA61234.1; -.
DR PIR; I40645; I40645.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR006025; Zn_MTpeptdse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE NEG.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.

FT	INIT MET	0	0	BY SIMILARITY.
FT	CHAIN	1	447	BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT	CHAIN	448	1295	BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT	METAL	222	222	ZINC (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	223	223	BY SIMILARITY.
FT	METAL	226	226	ZINC (CATALYTIC) (BY SIMILARITY).
FT	DISULFID	429	453	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	1234	1279	BY SIMILARITY.
FT	TRANSMEM	626	646	POTENTIAL.
FT	TRANSMEM	655	675	POTENTIAL.
SQ	SEQUENCE	1295 AA; 149279 MW; 5DA04A13D98D6372 CRC64;		

Query Match 50.0%; Score 56; DB 1; Length 1295;
Best Local Similarity 50.0%; Pred. No. 0.39;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14
: ||: |||::||
Db 937 YENFSTSFWIKIPK 950

RESULT 10

VNSS_INSVN

ID VNSS_INSVN STANDARD; PRT; 449 AA.

AC Q01811;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-JUN-1994 (Rel. 29, Last annotation update)

DE Nonstructural protein NS-S.

GN NSS.

OS Impatiens necrotic spot virus (strain NL-07) (INSV).

OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Tospovirus.

OX NCBI_TaxID=31622;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92331780; PubMed=1385787;

RA de Haan P., de Avila A.C., Kormelink R., Westerbroek A.,

RA Gielen J.J., Peters D., Goldbach R.;

RT "The nucleotide sequence of the S RNA of Impatiens necrotic spot

RT virus, a novel tospovirus.";

RL FEBS Lett. 306:27-32(1992).

CC -!- FUNCTION: FORMS FILAMENTOUS INCLUSION BODIES.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X66972; CAA47382.1; -.

DR PIR; S23158; S23158.

DR InterPro; IPR004915; Bunya_NS-S_2.

DR Pfam; PF03231; Bunya_NS-S_2; 1.

KW Nonstructural protein.

SQ SEQUENCE 449 AA; 51197 MW; C46AC1372B114CA5 CRC64;

Query Match 46.4%; Score 52; DB 1; Length 449;
Best Local Similarity 45.0%; Pred. No. 0.58;
Matches 9; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

Qy 2 NN--FTVSFWLRVPKVSASH 19
|| | : | | : | | : | :
Db 240 NNKPFKISLWMRIPKIMKSN 259

RESULT 11

VNSS_TSWV1

ID VNSS_TSWV1 STANDARD; PRT; 464 AA.
AC P26002;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Nonstructural protein NS-S.
GN NSS.
OS Tomato spotted wilt virus (strain Brazilian Br-01) (TSWV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Tospovirus.
OX NCBI_TaxID=36413;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90264829; PubMed=1693160;
RA de Haan P., Wagemakers L., Peters D., Goldbach R.;
RT "The S RNA segment of tomato spotted wilt virus has an ambisense
RT character.";
RL J. Gen. Virol. 71:1001-1007(1990).
CC -!- FUNCTION: FORMS FILAMENTOUS INCLUSION BODIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00645; BAA00540.1; -.
DR PIR; JQ0547; MNVUWC.
DR InterPro; IPR004915; Bunya_NS-S_2.
DR Pfam; PF03231; Bunya_NS-S_2; 1.
KW Nonstructural protein.
SQ SEQUENCE 464 AA; 52448 MW; AE5519179F9EF377 CRC64;

Query Match 46.4%; Score 52; DB 1; Length 464;
Best Local Similarity 64.3%; Pred. No. 0.6;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 NNFTVSVFWLRVPKV 15
::| :| |||||
Db 239 SHFKLSLWLRVPKV 252

RESULT 12 VNSS_TSWVL

ID VNSS_TSWVL STANDARD; PRT; 467 AA.
AC P26003;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Nonstructural protein NS-S.
GN NSS.
OS Tomato spotted wilt virus (strain Bulgarian L3) (TSWV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Tospovirus.
OX NCBI_TaxID=36415;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=91132150; PubMed=1993884;
 RA Maiss E., Ivanova L., Breyel E., Adam G.;
 RT "Cloning and sequencing of the S RNA from a Bulgarian isolate of
 RT tomato spotted wilt virus.";
 RL J. Gen. Virol. 72:461-464(1991).
 CC -!- FUNCTION: FORMS FILAMENTOUS INCLUSION BODIES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D13926; BAA03024.1; -.
 DR PIR; JQ0954; MNVUW1.
 DR InterPro; IPR004915; Bunya_NS-S_2.
 DR Pfam; PF03231; Bunya_NS-S_2; 1.
 KW Nonstructural protein.
 SQ SEQUENCE 467 AA; 52413 MW; 38E5CA4E802DB6DC CRC64;

Query Match 46.4%; Score 52; DB 1; Length 467;
 Best Local Similarity 64.3%; Pred. No. 0.6;
 Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 NNFTVSFWLRVPKV 15
 : : | : | | | | | |
 Db 243 SHFKLSLWLRVPKV 256

RESULT 13

BXCN_CLOBO
 ID BXCN_CLOBO STANDARD; PRT; 1196 AA.
 AC P46081;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Botulinum neurotoxin type C1, nontoxic component.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Type C Stockholm / C-ST;
 RX MEDLINE=92231894; PubMed=1567404;
 RA Tsuzuki K., Kimura K., Fujii N., Yokosawa N., Oguma K.;
 RT "The complete nucleotide sequence of the gene coding for the
 RT nontoxic-nonhemagglutinin component of Clostridium botulinum type C
 RT progenitor toxin.";
 RL Biochem. Biophys. Res. Commun. 183:1273-1279(1992).
 CC -!- FUNCTION: THE NONTOXIC COMPONENT IS NECESSARY TO MAINTAIN
 CC TOXICITY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X62389; CAA44262.1; -.
DR InterPro; IPR000395; Bontoxilysin.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
KW Neurotoxin.
SQ SEQUENCE 1196 AA; 138740 MW; 4BD5956274D7F9C3 CRC64;

Query Match 45.5%; Score 51; DB 1; Length 1196;
Best Local Similarity 80.0%; Pred. No. 2.4;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNFTVSFWLR 11
|||: |||
Db 899 NNFTICFWLR 908

RESULT 14

BXC1_CLOBO

ID BXC1_CLOBO STANDARD; PRT; 1290 AA.
AC P18640;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BoNT/C1)
DE (Bontoxilysin C1).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90370487; PubMed=2204031;
RA Hauser D., Eklund M.W., Kurazono H., Binz T., Niemann H., Gill D.M.,
RA Boquet P., Popoff M.R.;
RT "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";
RL Nucleic Acids Res. 18:4924-4924(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Type C Stockholm / C-ST;
RX MEDLINE=91024998; PubMed=2222445;
RA Kimura K., Fujii N., Tsuzuki K., Murakami T., Indoh T.,
RA Yokosawa N., Takeshi K., Syuto B., Oguma K.;
RT "The complete nucleotide sequence of the gene coding for botulinum
RT type C1 toxin in the C-ST phage genome.";
RL Biochem. Biophys. Res. Commun. 171:1304-1311(1990).
RN [3]
RP SEQUENCE OF 2-25.
RC STRAIN=Type C Stockholm / C-ST;
RX MEDLINE=88153072; PubMed=2450068;
RA Tsuzuki K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K.,

RA Oguma K.;
 RT "Establishment of a monoclonal antibody recognizing an antigenic site
 RT common to Clostridium botulinum type B, C1, D, and E toxins and
 RT tetanus toxin.";
 RL Infect. Immun. 56:898-902(1988).
 RN [4]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94038966; PubMed=7901002;
 RA Blasi J., Chapman E.R., Yamasaki S., Binz T., Niemann H., Jahn R.;
 RT "Botulinum neurotoxin C1 blocks neurotransmitter release by means of
 RT cleaving HPC-1/syntaxin.";
 RL EMBO J. 12:4821-4828(1993).

CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES SYNTAXIN.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -!- MISCELLANEOUS: BOTULINUM TYPE C1 NEUROTOXIN IS SYNTHESIZED BY C
 CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
 CC BACTERIOPHAGE.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X66433; CAA47060.1; -.
 DR EMBL; X72793; CAA51313.1; -.
 DR EMBL; X53751; CAA37780.1; -.
 DR EMBL; D90210; BAA14235.1; -.
 DR EMBL; X62389; CAA44263.1; -.
 DR HSSP; P10845; 3BTA.
 DR MEROPS; M27.002; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT INIT_MET 0 0

FT	CHAIN	1	448	BOTULINUM NEUROTOXIN C1, LIGHT-CHAIN.
FT	CHAIN	449	1290	BOTULINUM NEUROTOXIN C1, HEAVY-CHAIN.
FT	METAL	228	228	ZINC (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	229	229	BY SIMILARITY.
FT	METAL	232	232	ZINC (CATALYTIC) (BY SIMILARITY).
FT	DISULFID	436	452	INTERCHAIN (PROBABLE).
FT	CONFLICT	84	84	P -> T (IN REF. 2).
SQ	SEQUENCE	1290 AA; 148734 MW; 71FBE379F97129E8 CRC64;		

Query Match 42.9%; Score 48; DB 1; Length 1290;
 Best Local Similarity 42.9%; Pred. No. 8;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
 : :|::|||:|: |
 Db 934 YESFSISFWIRINK 947

RESULT 15

BXD_CLOBO

ID BXD_CLOBO STANDARD; PRT; 1276 AA.
 AC P19321;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BoNT/D)
 DE (Bontoxilysin D).
 GN BOTD.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BVD/-3;
 RX MEDLINE=91016853; PubMed=2216736;
 RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,
 RA Kozaki S., Krieglstein K., Henschen A., Gill D.M., Niemann H.;
 RT "Nucleotide sequence of the gene encoding Clostridium botulinum
 RT neurotoxin type D.";
 RL Nucleic Acids Res. 18:5556-5556(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CB16;
 RX MEDLINE=93042276; PubMed=1420572;
 RA Sunagawa H., Ohyama T., Watanabe T., Inoue K.;
 RT "The complete amino acid sequence of the Clostridium botulinum type D
 RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
 RT phage d-16 phi genome.";
 RL J. Vet. Med. Sci. 54:905-913(1992).
 RN [3]
 RP PARTIAL SEQUENCE.
 RC STRAIN=D-SA, and D-1873;
 RX MEDLINE=89339741; PubMed=2668193;
 RA Moriishi K., Syuto B., Kubo S., Oguma K.;
 RT "Molecular diversity of neurotoxins from Clostridium botulinum type D
 RT strains.";

RL Infect. Immun. 57:2886-2891(1989).
 RN [4]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94230352; PubMed=8175689;
 RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
 RA Roques B., Fykse E.M., Suedhof T.C., Jahn R., Niemann H.;
 RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
 RT F botulinal neurotoxins and tetanus toxin.";
 RL J. Biol. Chem. 269:12764-12772(1994).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 60-LYS-|-LEU-61 BOND OF
 CC SYNAPTOBREVINS-1 AND -2.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -!- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D
 CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
 CC BACTERIOPHAGE.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X54254; CAA38175.1; -.
 DR EMBL; S49407; AAB24244.1; -.
 DR PIR; S11455; S11455.
 DR HSSP; P10845; 3BTA.
 DR MEROPS; M27.002; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT CHAIN 1 442 BOTULINUM NEUROTOXIN D, LIGHT-CHAIN.
 FT CHAIN 443 1276 BOTULINUM NEUROTOXIN D, HEAVY-CHAIN.
 FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 230 230 BY SIMILARITY.
 FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).

FT	DISULFID	437	450	INTERCHAIN (PROBABLE).
FT	VARIANT	15	16	ND -> PV (IN STRAIN D-SA).
FT	VARIANT	17	18	ND -> LQ (IN STRAIN D-1873).
FT	VARIANT	452	452	K -> Q (IN STRAIN D-SA).
FT	VARIANT	457	457	R -> T (IN STRAIN D-SA).
FT	VARIANT	457	457	R -> F (IN STRAIN D-1873).
FT	VARIANT	462	462	A -> D (IN STRAIN D-1873).
FT	VARIANT	489	489	K -> N (IN STRAIN CB16).
FT	VARIANT	644	644	N -> K (IN STRAIN CB16).
FT	VARIANT	1122	1122	Q -> R (IN STRAIN CB16).
SQ	SEQUENCE	1276	AA; 146871	MW; C1EC50F46C8233E2 CRC64;

Query Match 42.4%; Score 47.5; DB 1; Length 1276;
 Best Local Similarity 40.9%; Pred. No. 9.6;
 Matches 9; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

Qy	1 FNNFTVSFWLRVPK-VSASHLE 21
	: : ::: ::
Db	928 YENSSVSFWIKISKDLTNSHNE 949

Search completed: December 13, 2003, 13:35:07
 Job time : 1.71795 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 2.89744 Seconds
(without alignments)
1870.310 Million cell updates/sec

Title: US-09-785-215-6
Perfect score: 112
Sequence: 1 FNNFTVSVFWLRVPKVSASHLE 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match	Length	DB	ID	Description

1	112	100.0	451	2	Q9LA13	Q9la13 clostridium
2	112	100.0	1310	2	Q93N27	Q93n27 clostridium
3	62	55.4	1268	2	Q45851	Q45851 clostridium
4	61	54.5	1278	2	Q57236	Q57236 clostridium
5	58	51.8	361	2	Q45848	Q45848 clostridium
6	58	51.8	361	2	Q45846	Q45846 clostridium
7	58	51.8	441	2	Q9X708	Q9x708 clostridium
8	58	51.8	1291	2	Q9ZAJ8	Q9zaj8 clostridium
9	58	51.8	1291	2	Q93G71	Q93g71 clostridium
10	58	51.8	1291	2	Q933K0	Q933k0 clostridium
11	58	51.8	1291	2	Q08077	Q08077 clostridium
12	58	51.8	1291	2	Q8GR96	Q8gr96 clostridium
13	56	50.0	367	2	Q45861	Q45861 clostridium
14	56	50.0	367	2	Q45862	Q45862 clostridium
15	56	50.0	1251	2	Q9K395	Q9k395 clostridium
16	56	50.0	1252	2	Q8KZM3	Q8kzm3 clostridium
17	56	50.0	1255	2	Q9FAR6	Q9far6 clostridium
18	55	49.1	1280	2	Q9ZAJ5	Q9zaj5 clostridium
19	52	46.4	466	12	Q8JXJ9	Q8jxj9 tomato spot
20	52	46.4	466	12	Q8JXK0	Q8jxk0 tomato spot
21	52	46.4	467	12	Q8JXK2	Q8jxk2 tomato spot
22	52	46.4	467	12	O37369	O37369 tomato spot
23	52	46.4	467	12	O37367	O37367 tomato spot
24	52	46.4	467	12	Q8JVL0	Q8jvl0 tomato spot
25	52	46.4	467	12	Q8JXK4	Q8jxk4 tomato spot
26	51	45.5	467	12	Q88900	Q88900 tospovirus.
27	51	45.5	1196	2	Q45916	Q45916 clostridium
28	51	45.5	1196	2	Q53550	Q53550 clostridium
29	51	45.5	1196	2	Q9LBR2	Q9lbr2 clostridium
30	51	45.5	1196	2	Q9LBS8	Q9lbs8 clostridium
31	51	45.5	1196	2	Q93HT4	Q93ht4 clostridium
32	51	45.5	1196	9	Q9ZX77	Q9zx77 clostridium
33	51	45.5	1196	9	Q38197	Q38197 clostridium
34	50	44.6	503	10	Q948E8	Q948e8 oryza sativ
35	50	44.6	503	10	Q8H892	Q8h892 oryza sativ
36	49	43.8	262	10	Q8GZE7	Q8gze7 pyrocystis
37	49	43.8	276	5	Q9TZK5	Q9tzk5 caenorhabdi
38	48	42.9	504	10	Q9M1N3	Q9m1n3 arabidopsis
39	48	42.9	1285	2	Q45967	Q45967 clostridium
40	48	42.9	1285	2	Q9LBR1	Q9lbr1 clostridium
41	48	42.9	1291	2	Q93HT3	Q93ht3 clostridium
42	47.5	42.4	1275	12	Q9QTG7	Q9qtg7 clostridium
43	47.5	42.4	1280	2	Q45849	Q45849 clostridium
44	47.5	42.4	1280	2	Q9LBS7	Q9lbs7 clostridium
45	47	42.0	345	5	O62165	O62165 caenorhabdi

ALIGNMENTS

RESULT 1

Q9LA13

ID Q9LA13 PRELIMINARY; PRT; 451 AA.

AC Q9LA13;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Tetanus toxin (Fragment).
 OS Clostridium tetani.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=20886;
 RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;
 RT "Fragment C of Tetanus Toxin."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF154828; AAF73267.1; -.
 DR HSSP; P04958; 1A8D.
 DR InterPro; IPR001064; Crystallin.
 DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
 FT NON_TER 1 1
 SQ SEQUENCE 451 AA; 51823 MW; 69A8C5F030E6CD8E CRC64;

Query Match 100.0%; Score 112; DB 2; Length 451;
 Best Local Similarity 100.0%; Pred. No. 3.3e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
 |||||
 Db 83 FNNFTVSFWLRVPKVSASHLE 103

RESULT 2

Q93N27

ID Q93N27 PRELIMINARY; PRT; 1310 AA.
 AC Q93N27;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Tetanus toxin (Fragment).
 OS Clostridium tetani.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shumin Z., Dianliang L.;
 RT "Cloning and sequence analysis of tetanus toxin gene."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF389424; AAK72964.2; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR001064; Crystallin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 FT NON_TER 1 1
 FT NON_TER 1310 1310
 SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
 |||||
Db 948 FNNFTVSFWLRVPKVSASHLE 968

Query Match 55.4%; Score 62; DB 2; Length 1268;
Best Local Similarity 64.3%; Pred. No. 0.19;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

ID   Q57236          PRELIMINARY;          PRT;   1278 AA.
AC   Q57236; Q45863;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Botulinum neurotoxin type F (BONT/F protein).

```

GN BONT/F.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCTC 10281;
 RA Hutson R.A., Collins M.D.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE OF 635-1000 FROM N.A.
 RC STRAIN=NCTC 1028;
 RX MEDLINE=94013372; PubMed=8408542;
 RA Campbell K., East A.K., Collins M.D.;
 RT "Gene probes for identification of the botulinal neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F.";
 RL J. Clin. Microbiol. 31:2255-2262(1993).
 RN [4]
 RP SEQUENCE OF 1-27 FROM N.A.
 RC STRAIN=LANGELAND;
 RX MEDLINE=98404102; PubMed=9732534;
 RA East A.K., Bhandari M., Hielm S., Collins M.D.;
 RT "Analysis of the botulinum neurotoxin type F gene clusters in
 RT proteolytic and nonproteolytic Clostridium botulinum and Clostridium
 RT barati.";
 RL Curr. Microbiol. 37:262-268(1998).
 DR EMBL; X81714; CAA57358.1; -.
 DR EMBL; L35496; AAA23210.1; -.
 DR EMBL; X70821; CAA50152.1; -.
 DR EMBL; X99064; CAA67512.1; -.
 DR HSSP; P10845; 3BTA.
 DR MEROPS; M27.002; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin.
 SQ SEQUENCE 1278 AA; 147073 MW; A1BE1318431D6918 CRC64;

Query Match 54.5%; Score 61; DB 2; Length 1278;
 Best Local Similarity 57.1%; Pred. No. 0.29;
 Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
 : ||::|||:|:|
 Db 931 YQNFSISFWVRIPK 944

RESULT 5
 Q45848

ID Q45848 PRELIMINARY; PRT; 361 AA.
AC Q45848;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Botulinum neurotoxin type B (Fragment).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type B;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinal neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL; X70819; CAA50150.1; -.
DR HSSP; P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1 1
FT NON_TER 361 361
SQ SEQUENCE 361 AA; 42131 MW; A2E0FFFC81F9533D CRC64;

Query Match 51.8%; Score 58; DB 2; Length 361;
Best Local Similarity 64.3%; Pred. No. 0.25;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
| :|:||||:|:|
Db 290 FLDFSVSFWIRIPK 303

RESULT 6

Q45846
ID Q45846 PRELIMINARY; PRT; 361 AA.
AC Q45846;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Botulinum neurotoxin type B (Fragment).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type B;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinal neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL; X70814; CAA50145.1; -.

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Bont protein.
 GN BONT.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 3281;
 RX MEDLINE=98440323; PubMed=9767710;
 RA Santos-Buelga J., Collins M.D., East A.K.;
 RT "Characterization of the genes encoding the Botulinum neurotoxin
 RT complex in a strain of clostridium botulinum producing type B & F
 RT neurotoxins.";
 RL Curr. Microbiol. 37:312-318(1998).
 DR EMBL; Y13630; CAA73968.1; -.
 DR HSSP; P10845; 3BTA.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 SQ SEQUENCE 1291 AA; 150840 MW; E4D3B0E46AB2E735 CRC64;

Query Match 51.8%; Score 58; DB 2; Length 1291;
 Best Local Similarity 64.3%; Pred. No. 0.92;
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVVK 14
 | : | : | | | : | : | |
 Db 923 FLDFSVSFWIRIPK 936

RESULT 9

Q93G71

ID Q93G71 PRELIMINARY; PRT; 1291 AA.
 AC Q93G71;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Neurotoxin type B.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1436;
 RA Kirma N., Ferreira J.L., Baumstark B.R.;
 RT "Characterization of six type A strains of Clostridium botulinum that
 RT contain type B toxin gene sequences.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF295926; AAK97132.1; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_MTpeptdse.

DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1291 AA; 150824 MW; D7CA07BAE2EB8CD2 CRC64;

Query Match 51.8%; Score 58; DB 2; Length 1291;
Best Local Similarity 64.3%; Pred. No. 0.92;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
| : : : : : : : : : :
Db 923 FLDFSVSFWIRIPK 936

RESULT 10

Q933K0

ID Q933K0 PRELIMINARY; PRT; 1291 AA.
AC Q933K0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type B cryptic neurotoxin.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=593, and 588;
RA Kirma N., Ferreira J.L., Baumstark B.R.;
RT "Characterization of six type A strains of Clostridium botulinum that
RT contain type B toxin gene sequences.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF300466; AAL11499.1; -.
DR EMBL; AF300465; AAL11498.1; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR006025; Zn_MTpeptdse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin.
SQ SEQUENCE 1291 AA; 150843 MW; 7AC1737B0FA5A151 CRC64;

Query Match 51.8%; Score 58; DB 2; Length 1291;
Best Local Similarity 64.3%; Pred. No. 0.92;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
| : : : : : : : : : :
Db 923 FLDFSVSFWIRIPK 936

RESULT 11

Q08077

ID Q08077 PRELIMINARY; PRT; 1291 AA.

AC Q08077;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE BoNT/B.
 GN BONT/B.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Eklund 17B ATCC25765;
 RX MEDLINE=94122659; PubMed=7764370;
 RA Hutson R.A., Collins M.D., East A.K., Thompson D.E.;
 RT "Nucleotide sequence of the gene coding for non-proteolytic
 RT Clostridium botulinum type B neurotoxin: comparison with other
 RT clostridial neurotoxins."
 RL Curr. Microbiol. 28:101-110(1994).
 DR EMBL; X71343; CAA50482.1; -.
 DR HSSP; P10845; 3BTA.
 DR MEROPS; M27.002; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 SQ SEQUENCE 1291 AA; 150513 MW; 71BCAFE23D69FAAA CRC64;

Query Match 51.8%; Score 58; DB 2; Length 1291;
 Best Local Similarity 64.3%; Pred. No. 0.92;
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSVFWLRVPK 14
 | :|:||||:|:|
 Db 923 FLDFSVSFWIRIPK 936

RESULT 12

Q8GR96

ID Q8GR96 PRELIMINARY; PRT; 1291 AA.
 AC Q8GR96;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Neurotoxin.
 GN BONTB.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ihara H., Kohda T., Morimoto F., Tsukamoto K., Karasawa T.,
 RA Nakamura S., Mukamoto M., Kozaki S.;
 RT "Clostridium botulinum type B neurotoxin associated with infant

RT botulism.";
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB084152; BAC22064.1; -.
SQ SEQUENCE 1291 AA; 150574 MW; 0227CAEF4F58504D CRC64;

Query Match 51.8%; Score 58; DB 2; Length 1291;
Best Local Similarity 64.3%; Pred. No. 0.92;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
| :|:|||||:|:
Db 923 FLDFSVSFWIRIPK 936

RESULT 13

Q45861

ID Q45861 PRELIMINARY; PRT; 367 AA.
AC Q45861;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Botulinum neurotoxin type E (Fragment).
GN BONT/E.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type E;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinal neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL; X70818; CAA50149.1; -.
DR HSSP; P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1 1
FT NON_TER 367 367
SQ SEQUENCE 367 AA; 42902 MW; 346A610C2FF70262 CRC64;

Query Match 50.0%; Score 56; DB 2; Length 367;
Best Local Similarity 53.8%; Pred. No. 0.54;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVP 13
: ||:||||:|:
Db 297 YKNFSISFWVRIP 309

RESULT 14

Q45862

ID Q45862 PRELIMINARY; PRT; 367 AA.
AC Q45862;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE Botulinum neurotoxin type E (Fragment).
 GN BONT/E.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=type E;
 RX MEDLINE=94013372; PubMed=8408542;
 RA Campbell K., East A.K., Collins M.D.;
 RT "Gene probes for identification of the botulinal neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F.";
 RL J. Clin. Microbiol. 31:2255-2262(1993).
 DR EMBL; X70815; CAA50146.1; -.
 DR HSSP; P10845; 3BTA.
 KW Neurotoxin.
 FT NON_TER 1 1
 FT NON_TER 367 367
 SQ SEQUENCE 367 AA; 42854 MW; 0810595B3A865570 CRC64;

 Query Match 50.0%; Score 56; DB 2; Length 367;
 Best Local Similarity 53.8%; Pred. No. 0.54;
 Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

 Qy 1 FNNFTVSFWLRVP 13
 : ||::|||:|:
 Db 297 YKNFSISFWVRIP 309

RESULT 15
 Q9K395
 ID Q9K395 PRELIMINARY; PRT; 1251 AA.
 AC Q9K395;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Type E botulinum toxin.
 GN BONT/E.
 OS Clostridium butyricum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1492;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL 095;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (LCL 095) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL 155;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Gyobu Y., Yamakawa K.,
 RA Kato H., Nakamura S., Karasawa T.;
 RT "C. butyricum (LCL 155) gene for type E botulinum toxin.";

RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KZ 1899;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1899) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KZ 1897;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1897) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KZ 1898;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1898) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KZ 1886;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1886) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KZ 1887;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1887) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KZ 1889;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1889) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KZ 1890;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1890) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [10]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KZ 1891;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1891) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.

RN [11]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL 063;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (LCL 063) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AB037714; BAB03522.1; -.
 DR EMBL; AB037704; BAB03512.1; -.
 DR EMBL; AB037705; BAB03513.1; -.
 DR EMBL; AB037706; BAB03514.1; -.
 DR EMBL; AB037707; BAB03515.1; -.
 DR EMBL; AB037708; BAB03516.1; -.
 DR EMBL; AB037709; BAB03517.1; -.
 DR EMBL; AB037710; BAB03518.1; -.
 DR EMBL; AB037711; BAB03519.1; -.
 DR EMBL; AB037712; BAB03520.1; -.
 DR EMBL; AB037713; BAB03521.1; -.
 DR HSSP; P10845; 3BTA.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR006025; Zn_MTpeptdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 SQ SEQUENCE 1251 AA; 143751 MW; 2021F4E427070296 CRC64;

Query Match 50.0%; Score 56; DB 2; Length 1251;
 Best Local Similarity 53.8%; Pred. No. 1.9;
 Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVP 13
 : ||::|||:|:
 Db 912 YKNFSISFWVRIP 924

Search completed: December 13, 2003, 13:29:53
 Job time : 5.89744 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 1.79365 Seconds
(without alignments)
1870.310 Million cell updates/sec

Title: US-09-785-215-19
Perfect score: 64
Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	%	No.	Score	Match	Length	DB	ID	Description

1	43	67.2	370	16	Q8R8U0	Q8r8u0 thermoanaer
2	41	64.1	832	16	Q8XYF4	Q8xyf4 ralstonia s
3	40	62.5	261	2	Q8GB17	Q8gb17 proteus sp.
4	40	62.5	275	16	Q98CR0	Q98cr0 rhizobium l
5	40	62.5	422	2	Q9ETG0	Q9etg0 corynebacte
6	39	60.9	297	16	Q8FLA6	Q8fla6 escherichia
7	39	60.9	331	5	Q8STE0	Q8ste0 encephalito
8	39	60.9	612	5	Q8STA9	Q8sta9 encephalito
9	39	60.9	758	16	Q8DFQ8	Q8dfq8 vibrio vuln
10	38	59.4	96	2	Q93DU6	Q93du6 unidentifie
11	38	59.4	352	16	Q8YBT0	Q8ybt0 brucella me
12	38	59.4	352	16	Q8FWJ3	Q8fwj3 brucella su
13	38	59.4	438	12	O39493	O39493 bovine herp
14	38	59.4	639	16	Q92WJ9	Q92wj9 rhizobium m
15	38	59.4	698	10	Q94HP0	Q94hp0 oryza sativ
16	38	59.4	758	16	Q9KQC3	Q9kqc3 vibrio chol
17	38	59.4	826	16	Q8EIJ1	Q8eij1 shewanella
18	38	59.4	3037	12	Q68749	Q68749 hepatitis c
19	37	57.8	288	16	Q989E2	Q989e2 rhizobium l
20	37	57.8	309	16	O69692	O69692 mycobacteri
21	37	57.8	329	5	Q9VFY8	Q9vfy8 drosophila
22	37	57.8	368	16	Q9A7N1	Q9a7n1 caulobacter
23	37	57.8	391	5	Q9NKE6	Q9nke6 drosophila
24	37	57.8	490	3	Q8J0I9	Q8j0i9 trichoderma
25	37	57.8	612	16	Q987K9	Q987k9 rhizobium l
26	37	57.8	663	3	Q96X20	Q96x20 laccaria fr
27	37	57.8	678	5	Q9I7L8	Q9i7l8 drosophila
28	37	57.8	712	16	Q8YZ31	Q8yz31 anabaena sp
29	37	57.8	907	5	Q8IF42	Q8if42 trypanosoma
30	37	57.8	1205	3	O59946	O59946 emericella
31	37	57.8	1470	5	Q9VF24	Q9vf24 drosophila
32	37	57.8	3032	12	Q99IB2	Q99ib2 hepatitis c
33	37	57.8	3032	12	Q99IB4	Q99ib4 hepatitis c
34	37	57.8	3032	12	Q9Q9A9	Q9q9a9 hepatitis c
35	37	57.8	3033	12	Q9IZA5	Q9iza5 hepatitis c
36	37	57.8	3033	12	Q99IB6	Q99ib6 hepatitis c
37	37	57.8	3033	12	Q9Q9A7	Q9q9a7 hepatitis c
38	37	57.8	3033	12	Q9IZA3	Q9iza3 hepatitis c
39	37	57.8	3033	12	Q9IZA4	Q9iza4 hepatitis c
40	37	57.8	3033	12	Q9Q9B0	Q9q9b0 hepatitis c
41	37	57.8	3033	12	Q9IZA6	Q9iza6 hepatitis c
42	37	57.8	3033	12	Q99IB3	Q99ib3 hepatitis c
43	36	56.2	156	5	O60959	O60959 leishmania
44	36	56.2	176	17	Q8ZXN0	Q8zxn0 pyrobaculum
45	36	56.2	205	16	Q8PMF5	Q8pmf5 xanthomonas

ALIGNMENTS

RESULT 1

Q8R8U0

ID Q8R8U0 PRELIMINARY; PRT; 370 AA.
AC Q8R8U0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Kef-type K+ transport systems, membrane components.
 GN KEFB OR TTE1901.
 OS Thermoanaerobacter tengcongensis.
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.
 OX NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4 / JCM 11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of T. tengcongensis genome.";
 RL Genome Res. 12:689-700(2002).
 DR EMBL; AE013140; AAM25084.1; --
 DR InterPro; IPR006153; Na_H_porter.
 DR Pfam; PF00999; Na_H_Exchanger; 1.
 KW Complete proteome.
 SQ SEQUENCE 370 AA; 40695 MW; 489870521F84215A CRC64;

Query Match 67.2%; Score 43; DB 16; Length 370;
 Best Local Similarity 80.0%; Pred. No. 15;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAA 12
 ||| |||:|
 Db 101 FVAGWTLQAA 110

RESULT 2

Q8XYF4

ID Q8XYF4 PRELIMINARY; PRT; 832 AA.
 AC Q8XYF4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein RSc1804.
 GN RSC1804 OR RS04224.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Ralstoniaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GMI1000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choisine N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:497-502(2002).
 DR EMBL; AL646066; CAD15506.1; --
 DR InterPro; IPR004568; Pantethn_trn.

DR InterPro; IPR000408; Reg_chr_condens.
 DR InterPro; IPR000051; SAM_bind.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF01648; ACPS; 1.
 DR Pfam; PF00975; Thioesterase; 1.
 DR PROSITE; PS00626; RCC1_2; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 832 AA; 89947 MW; 8F679F2A62685FCA CRC64;

Query Match 64.1%; Score 41; DB 16; Length 832;
 Best Local Similarity 63.6%; Pred. No. 78;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAA 12
 :|:||||:|
 Db 751 RFLAAWSLKEA 761

RESULT 3

Q8GB17

ID Q8GB17 PRELIMINARY; PRT; 261 AA.
 AC Q8GB17;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Crotonobetainyl-CoA-hydratase.
 GN CAID.
 OS Proteus sp. LE138.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Proteus.
 OX NCBI_TaxID=217617;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Engemann C., Elssner T., Pfeifer S., Krumbholz C., Maier T.,
 RA Kleber H.P.;
 RT "Cai locus and corresponding enzymes of Proteus sp.";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AJ508908; CAD48582.1; -.
 SQ SEQUENCE 261 AA; 28184 MW; 87BF6AE0A553FE43 CRC64;

Query Match 62.5%; Score 40; DB 2; Length 261;
 Best Local Similarity 72.7%; Pred. No. 36;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13
 | | | | | | |
 Db 60 FCAGWDLKAAA 70

RESULT 4

Q98CR0

ID Q98CR0 PRELIMINARY; PRT; 275 AA.
 AC Q98CR0;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Enoyl-CoA hydratase (EC 4.2.1.17).

GN MLR5044.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP003005; BAB51561.1; -.
 DR InterPro; IPR001753; EnCoA_hydrtse.
 DR Pfam; PF00378; ECH; 1.
 DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
 KW Lyase; Complete proteome.
 SQ SEQUENCE 275 AA; 30148 MW; 9C52AE72609AC7D3 CRC64;

Query Match 62.5%; Score 40; DB 16; Length 275;
 Best Local Similarity 72.7%; Pred. No. 38;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13
 | | | | |
 Db 75 FCAGWDLKAAA 85

RESULT 5

Q9ETG0

ID Q9ETG0 PRELIMINARY; PRT; 422 AA.
 AC Q9ETG0;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical 42.2 kDa protein.
 GN ORF49.
 OS Corynebacterium equi (Rhodococcus equi).
 OG Plasmid pREAT701 (p33701), and Plasmid virulence.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Nocardiaceae; Rhodococcus.
 OX NCBI_TaxID=43767;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC33701; PLASMID=pREAT701 (p33701);
 RA Takai S., Sekizaki T., Kakuda T., Nakamura M., Suzuki K., Ogino N.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC33701, and 103; PLASMID=pREAT701 (p33701), and virulence;
 RX PubMed=11083803;
 RA Takai S., Hines S.A., Sekizaki T., Nicholson V.M., Alperin D.A.,

RA Osaki M., Takamatsu D., Nakamura M., Suzuki K., Ogino N., Kakuda T.,
 RA Dan H., Prescott J.F.;
 RT "DNA sequence and comparison of virulence plasmids from Rhodococcus
 RT equi ATCC 33701 and 103.";
 RL Infect. Immun. 68:6840-6847(2000).
 DR EMBL; AP001204; BAB16658.1; -.
 DR EMBL; AF116907; AAG21752.1; -.
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 422 AA; 42203 MW; D908E81B9151558B CRC64;

Query Match 62.5%; Score 40; DB 2; Length 422;
 Best Local Similarity 66.7%; Pred. No. 59;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KFVAAWTLKAAA 13
 ||||:| |||
 Db 329 KFVASWLAAAAA 340

RESULT 6

Q8FLA6

ID Q8FLA6 PRELIMINARY; PRT; 297 AA.
 AC Q8FLA6;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Carnitiny-CoA dehydratase (EC 4.2.1.-).
 GN CAID OR C0045.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 DR EMBL; AE016755; AAN78543.1; -.
 KW Lyase; Complete proteome.
 SQ SEQUENCE 297 AA; 32285 MW; 9F568751F97EB885 CRC64;

Query Match 60.9%; Score 39; DB 16; Length 297;
 Best Local Similarity 72.7%; Pred. No. 62;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FVAAWTLKAAA 13
 |||||
 Db 96 FSAGWDLKAAA 106

RESULT 7

Q8STE0

ID Q8STE0 PRELIMINARY; PRT; 331 AA.
AC Q8STE0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein ECU01_0080 (Hypothetical protein ECU01_1530)
DE (Hypothetical protein ECU02_1560) (Hypothetical protein ECU04_0090)
DE (Hypothetical protein ECU08_0010) (Hypothetical protein
DE ECU08_2090).
GN ECU01_0080 OR ECU01_1530 OR ECU02_1560 OR ECU04_0090 OR ECU08_0010 OR
GN ECU08_2090.
OS Encephalitozoon cuniculi.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL391737; CAD24880.1; -.
DR EMBL; AL391737; CAD25024.1; -.
DR EMBL; AL590442; CAD25185.1; -.
DR EMBL; AL590444; CAD25196.1; -.
DR EMBL; AL590448; CAD26306.1; -.
DR EMBL; AL590448; CAD26511.1; -.
KW Hypothetical protein.
SQ SEQUENCE 331 AA; 37888 MW; 5F54620501D4DF54 CRC64;

Query Match 60.9%; Score 39; DB 5; Length 331;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13
: ||||| |||
Db 3 RIYAAWTLVAAA 14

RESULT 8

Q8STA9

ID Q8STA9 PRELIMINARY; PRT; 612 AA.

AC Q8STA9;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein ECU11_0050 (Hypothetical protein
 DE ECU05_1680).
 GN ECU11_0050 OR ECU05_1680.
 OS Encephalitozoon cuniculi.
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
 OX NCBI_TaxID=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RA Genoscope;
 RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RX MEDLINE=21576510; PubMed=11719806;
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
 RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
 RA Weissenbach J., Vivares C.P.;
 RT "Genome sequence and gene compaction of the eukaryote parasite
 RT Encephalitozoon cuniculi.";
 RL Nature 414:450-453(2001).
 DR EMBL; AL590450; CAD25915.1; -.
 DR EMBL; AL590445; CAD26688.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 612 AA; 70929 MW; 924DE330BA373593 CRC64;

Query Match 60.9%; Score 39; DB 5; Length 612;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KFVAAWTLKAAA 13
 : ||||| |||
 Db 3 RIYAAWTLVAAA 14

RESULT 9

Q8DFQ8

ID Q8DFQ8 PRELIMINARY; PRT; 758 AA.
 AC Q8DFQ8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Fe2+ transport system protein B.
 GN VV10148.
 OS Vibrio vulnificus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=672;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CMCP6;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,

RA Choy H.E.;
RT "Complete genome sequence of *Vibrio vulnificus* CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AE016797; AAO08686.1; -.
KW Complete proteome.
SQ SEQUENCE 758 AA; 82450 MW; 3ABE99077B86A57D CRC64;

Query Match 60.9%; Score 39; DB 16; Length 758;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAA 12
|:|:| |:|
Db 693 ARFIAVWTMALA 704

RESULT 10

Q93DU6

ID Q93DU6 PRELIMINARY; PRT; 96 AA.
AC Q93DU6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 10.6 kDa protein.
OS unidentified bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=2338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21536188; PubMed=11679351;
RA Stokes H.W., Holmes A.J., Nield B.S., Holley M.P., Nevalainen K.M.H.,
RA Mabbutt B.C., Gillings M.R.;
RT "Gene Cassette PCR: Sequence-Independent Recovery of Entire Genes from
RT Environmental DNA.";
RL Appl. Environ. Microbiol. 67:5240-5246(2001).
DR EMBL; AF378540; AAK56385.1; -.
KW Hypothetical protein.
SQ SEQUENCE 96 AA; 10602 MW; 80D289940D404593 CRC64;

Query Match 59.4%; Score 38; DB 2; Length 96;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAA 12
|:|:| |:|
Db 52 ARFIATWRLSSA 63

RESULT 11

Q8YBT0

ID Q8YBT0 PRELIMINARY; PRT; 352 AA.
AC Q8YBT0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative transmembrane protein.
GN BMEII0806.

OS *Brucella melitensis*.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; *Brucella*.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselkorn R., Kyrpides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT *Brucella melitensis*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL; AE009714; AAL54048.1; -.
 DR InterPro; IPR002549; UPF0118.
 DR Pfam; PF01594; UPF0118; 1.
 KW Complete proteome.
 SQ SEQUENCE 352 AA; 39062 MW; 6D7CDA913601125F CRC64;

Query Match 59.4%; Score 38; DB 16; Length 352;
 Best Local Similarity 63.6%; Pred. No. 1.1e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKA 11
 | | : | | | : | |
 Db 331 AMFIAAWSLLA 341

RESULT 12

Q8FWJ3

ID Q8FWJ3 PRELIMINARY; PRT; 352 AA.
 AC Q8FWJ3;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Membrane protein, putative.
 GN BRA0460.
 OS *Brucella suis*.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; *Brucella*.
 OX NCBI_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The *Brucella suis* genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).

DR EMBL; AE014544; AAN33654.1; -.
DR TIGR; BRA0460; -.
KW Complete proteome.
SQ SEQUENCE 352 AA; 39106 MW; AE6E4B8E3613829D CRC64;

Query Match 59.4%; Score 38; DB 16; Length 352;
Best Local Similarity 63.6%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKA 11
| | : | | | : | |
Db 331 AMFIAAWSLLA 341

RESULT 13

O39493

ID O39493 PRELIMINARY; PRT; 438 AA.
AC O39493;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glycoprotein M.
GN UL10.
OS Bovine herpesvirus type 1 (strain Jura).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Jura;
RX MEDLINE=95313343; PubMed=7793062;
RA Vlcek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
RA Letchworth G.J., Schwyzner M.;
RT "Nucleotide sequence analysis of a 30-kb region of the bovine
RT herpesvirus 1 genome which exhibits a colinear gene arrangement with
RT the UL21 to UL4 genes of herpes simplex virus.";
RL Virology 210:100-108(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Jura;
RA Schwyzner M., Paces V., Letchworth G.J., Misra V., Buhk H.J.,
RA Lowery D.E., Simard C., Bello L.J., Thiry E., Vlcek C.;
RT "Complete DNA sequence of bovine herpesvirus 1.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ004801; CAA06124.1; -.
DR InterPro; IPR000785; Herpes_glycop.
DR Pfam; PF01528; Herpes_glycop; 1.
DR PRINTS; PR00333; HSVINTEGRALMP.
SQ SEQUENCE 438 AA; 45517 MW; 4E7C7FA64FAAEFC7 CRC64;

Query Match 59.4%; Score 38; DB 12; Length 438;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VAAWTLKAA 12
: | | | | : | |
Db 141 LAAWTLQAA 149

RESULT 14

Q92WJ9

ID Q92WJ9 PRELIMINARY; PRT; 639 AA.
AC Q92WJ9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein RB0340.
GN RB0340 OR SMB20354.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymB (megaplasmid 2).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL603643; CAC48740.1; -.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 639 AA; 69534 MW; BD4335709C48FC5C CRC64;

Query Match 59.4%; Score 38; DB 16; Length 639;
Best Local Similarity 69.2%; Pred. No. 2e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
| | | | | | | |
Db 491 ALFGAAWTLAYAA 503

RESULT 15

Q94HP0

ID Q94HP0 PRELIMINARY; PRT; 698 AA.
AC Q94HP0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative retroelement.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nipponbare;
RA Spiegel L., de la Bastide M., Nascimento L., Kirchoff K., King L.,
RA Preston R., Vil M.D., Baker J., Bell M., Zutavern T., Santos L.,

RA Miller B., Kuit K., Rodriguez S., Cunnius D.M., Baliya V., Shah R.,
 RA Bahret A., Bal H., O'Shaughnessy A., Dedhia N., McCombie W.R.;
 RT "Genomic Sequence for Oryza sativa, Nipponbare strain, clone
 RT OSJNBa0089D15, from Chromosome 10, complete sequence.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nipponbare;
 RA McCombie W.R.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nipponbare;
 RA McCombie W.R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nipponbare;
 RA Palmer L.E., Spiegel L., de la Bastide M., Nascimento L., Kirchoff K.,
 RA King L., Preston R., Vil M.D., Baker J., Bell M., Zutavern T.,
 RA Santos L., Miller B., Kuit K., Rodriguez S., Cunnius D.M., Baliya V.,
 RA Shah R., Bahret A., Bal H., O'Shaughnessy A., Dedhia N.,
 RA McCombie W.R.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AC078944; AAK92597.1; -.
 DR Gramene; Q94HP0; -.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF00075; rnaseH; 1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00078; rvt; 1.
 KW RNA-directed DNA polymerase; Transferase.
 SQ SEQUENCE 698 AA; 78285 MW; 7ABD8BD52DDACC3A CRC64;

Query Match 59.4%; Score 38; DB 10; Length 698;
 Best Local Similarity 61.5%; Pred. No. 2.2e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13
 | ||| ||| |:
 Db 214 ADFVAEWTLAFAS 226

Search completed: December 13, 2003, 13:29:56
 Job time : 4.79365 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 0.444444 Seconds
(without alignments)
1375.531 Million cell updates/sec

Title: US-09-785-215-19
Perfect score: 64
Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query			ID	Description
		Match	Length	DB		
1	40	62.5	260	1	CAID_PROSL	Q8gb17 proteus sp.
2	39	60.9	146	1	YK81_MYCTU	Q10689 mycobacteri
3	39	60.9	260	1	CAID_ECO57	Q8xa35 escherichia
4	39	60.9	260	1	CAID_ECOL6	Q8fla6 escherichia
5	39	60.9	260	1	CAID_ECOLI	P31551 escherichia
6	39	60.9	260	1	CAID_SALTI	Q8z915 salmonella
7	39	60.9	260	1	CAID_SALTY	Q8zrx5 salmonella
8	39	60.9	260	1	CAID_SHIFL	P59395 shigella fl
9	38	59.4	155	1	LACB_MACGI	P11944 macropus gi
10	38	59.4	174	1	LACB_MACEU	Q29614 macropus eu
11	38	59.4	204	1	UBCX_PICPA	P49428 pichia past
12	38	59.4	411	1	VGLM_HSVBC	P52370 bovine herp
13	37	57.8	410	1	APGM_PYRAB	Q9v2m6 pyrococcus
14	37	57.8	411	1	APGM_PYRFU	P58814 pyrococcus
15	36	56.2	362	1	YAHA_ECOLI	P21514 escherichia
16	36	56.2	776	1	TFRI_CHICK	Q90997 gallus gall
17	35	54.7	128	1	U384_HSVMG	Q05102 marek's dis

18	35	54.7	550	1	PTR2_HUMAN	P49190	homo sapien
19	35	54.7	639	1	V70K_PLRVW	P11622	potato leaf
20	35	54.7	730	1	EF2_METAC	Q8trc3	methanosarc
21	35	54.7	730	1	EF2_METMA	Q8pur7	methanosarc
22	35	54.7	730	1	EF2_METTE	O93640	methanosarc
23	35	54.7	750	1	FOH1_HUMAN	Q04609	h glutamate
24	35	54.7	976	1	HMDH_GIBFU	Q12577	gibberella
25	35	54.7	1036	1	Y946_ARCFU	O29316	archaeoglob
26	35	54.7	1068	1	HMDH_ASPTE	Q9y7d2	aspergillus
27	34	53.1	83	1	HPIS_THET1	P80176	thermochrom
28	34	53.1	85	1	HPIS_THIRO	P00261	thiocapsa r
29	34	53.1	122	1	HPIS_CHRVI	P00260	chromatium
30	34	53.1	379	1	SGAA_METEX	P55819	methylobact
31	34	53.1	503	1	C726_ARATH	O65787	arabidopsis
32	34	53.1	574	1	KCN2_MOUSE	P58390	mus musculu
33	34	53.1	579	1	KCN2_HUMAN	Q9h2s1	homo sapien
34	34	53.1	580	1	KCN2_RAT	P70604	rattus norv
35	34	53.1	668	1	PGTB_SALTY	P37433	salmonella
36	34	53.1	739	1	CATA_MYCSM	Q59557	mycobacteri
37	34	53.1	3033	1	POLG_HCVJ6	P26660	h genome po
38	33	51.6	76	1	CXO2_CONTE	Q9xzk9	conus texti
39	33	51.6	141	1	LYSA_BPP2	P51769	bacterioph
40	33	51.6	188	1	UBCX_PICAN	O60015	pichia angu
41	33	51.6	214	1	DPOL_ADET1	Q88469	tupaia aden
42	33	51.6	237	1	Y751_TREPA	O83732	treponema p
43	33	51.6	291	1	LIP_THELA	O59952	thermomyces
44	33	51.6	381	1	TAL1_STRCO	O88018	streptomyce
45	33	51.6	412	1	APGM_PYRHO	O57742	pyrococcus

ALIGNMENTS

RESULT 1

CAID_PROSL

ID CAID_PROSL STANDARD; PRT; 260 AA.

AC Q8GB17;

DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Carnitiny1-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA hydratase).

GN CAID.

OS Proteus sp. (strain LE138).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Proteus.

OX NCBI_TaxID=217617;

RN [1]

RP SEQUENCE FROM N.A.

RA Engemann C., Elssner T., Pfeifer S., Krumbholz C., Maier T.,

RA Kleber H.-P.;

RT "Cai locus and corresponding enzymes of Proteus sp.";

RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.

RN [2]

RP SEQUENCE OF 1-20, AND SUBUNIT.

RX MEDLINE=21302209; PubMed=11409545;

RA Engemann C., Elssner T., Kleber H.-P.;

RT "Biotransformation of crotonobetaine to L(-)-carnitine in Proteus
 RT sp.";
 RL Arch. Microbiol. 175:353-359(2001).
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitiny-CoA
 CC to crotonobetainyl-CoA (By similarity).
 CC -!- CATALYTIC ACTIVITY: L-carnitiny-CoA = H(2)O + crotonobetainyl-
 CC CoA.
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to gamma-
 CC butyrobetaine); second step.
 CC -!- SUBUNIT: Homotrimer (Probable).
 CC -!- SIMILARITY: BELONGS TO THE ENOYL-COA HYDRATASE/ISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ508908; CAD48582.1; -.
 DR HAMAP; MF_01051; -; 1.
 DR Pfam; PF00378; ECH; 1.
 DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
 KW Lyase.
 FT INIT_MET 0 0
 FT ACT_SITE 110 110 BASE (BY SIMILARITY).
 FT ACT_SITE 130 130 ACID (BY SIMILARITY).
 SQ SEQUENCE 260 AA; 28053 MW; A6375AE0BD12CE43 CRC64;

Query Match 62.5%; Score 40; DB 1; Length 260;
 Best Local Similarity 72.7%; Pred. No. 5;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FVAAWTLKAAA 13
 | | | | |
 Db 59 FCAGWDLKAAA 69

RESULT 2

YK81_MYCTU

ID YK81_MYCTU STANDARD; PRT; 146 AA.

AC Q10689;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein Rv2081c.

GN RV2081C OR MT2143 OR MTCY49.20C.

OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37Rv;

RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z73966; CAA98193.1; -.
 DR EMBL; AE007063; AAK46425.1; ALT_INIT.
 DR PIR; D70766; D70766.
 DR TIGR; MT2143; -.
 DR TubercuList; Rv2081c; -.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 1 21 POTENTIAL.
 FT TRANSMEM 35 55 POTENTIAL.
 FT TRANSMEM 87 107 POTENTIAL.
 FT TRANSMEM 111 131 POTENTIAL.
 SQ SEQUENCE 146 AA; 14238 MW; BCA294986F8C9FD2 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 146;
 Best Local Similarity 63.6%; Pred. No. 4.4;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 FVAAWTLKAAA 13
 ||| || :||:
 Db 10 FVAIWTARAAS 20

RESULT 3
 CAID_ECO57
 ID CAID_ECO57 STANDARD; PRT; 260 AA.
 AC Q8XA35;
 DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Carnitiny-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA
 DE hydratase).
 GN CAID OR Z0042 OR ECS0039.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamosis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitiny-CoA
 CC to crotonobetainyl-CoA (By similarity).
 CC -!- CATALYTIC ACTIVITY: L-carnitiny-CoA = H(2)O + crotonobetainyl-
 CC CoA.
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to
 CC gamma-butyrobetaine); second step.
 CC -!- SIMILARITY: BELONGS TO THE ENOYL-COA HYDRATASE/ISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE005180; AAG54339.1; ALT_INIT.
 DR EMBL; AP002550; BAB33462.1; ALT_INIT.
 DR HAMAP; MF_01051; -, 1.
 DR InterPro; IPR001753; EnCoA_hydrase.
 DR Pfam; PF00378; ECH; 1.
 DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
 KW Lyase; Complete proteome.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT ACT_SITE 110 110 BASE (BY SIMILARITY).

FT ACT_SITE 130 130 ACID (BY SIMILARITY).
SQ SEQUENCE 260 AA; 28006 MW; 7E1E9CC3C9290ECF CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;
Best Local Similarity 72.7%; Pred. No. 7.6;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FVAAWTLKAAA 13
| | | | |
Db 59 FSAGWDLKAAA 69

RESULT 4

CAID_ECOL6

ID CAID_ECOL6 STANDARD; PRT; 260 AA.
AC Q8FLA6;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Carnitiny-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA
DE hydratase).
GN CAID OR C0045.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitiny-CoA
CC to crotonobetainyl-CoA (By similarity).
CC -!- CATALYTIC ACTIVITY: L-carnitiny-CoA = H(2)O + crotonobetainyl-
CC CoA.
CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to gamma-
CC butyrobetaine); second step.
CC -!- SIMILARITY: BELONGS TO THE ENOYL-COA HYDRATASE/ISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE016755; AAN78543.1; ALT_INIT.
DR HAMAP; MF_01051; -; 1.
DR Pfam; PF00378; ECH; 1.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.

KW Lyase; Complete proteome.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT ACT_SITE 110 110 BASE (BY SIMILARITY).
 FT ACT_SITE 130 130 ACID (BY SIMILARITY).
 SQ SEQUENCE 260 AA; 28045 MW; F33409DD2628B3B6 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;
 Best Local Similarity 72.7%; Pred. No. 7.6;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13
 | | | | |
 Db 59 FSAGWDLKAAA 69

RESULT 5

CAID_ECOLI

ID CAID_ECOLI STANDARD; PRT; 260 AA.
 AC P31551; P75623;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Carnitiny-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA
 DE hydratase).
 GN CAID OR B0036.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 RT the 0-2.4 min region.";
 RL Nucleic Acids Res. 20:3305-3308(1992).
 RN [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=O44:K74;
 RX MEDLINE=95115548; PubMed=7815937;
 RA Eichler K., Bourgis F., Buchet A., Kleber H.-P.,
 RA Mandrand-Berthelot M.-A.;
 RT "Molecular characterization of the cai operon necessary for carnitine
 RT metabolism in Escherichia coli.";
 RL Mol. Microbiol. 13:775-786(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).

RN [4]
 RP SEQUENCE OF 1-10, AND CHARACTERIZATION.
 RC STRAIN=O44:K74;
 RX MEDLINE=21435666; PubMed=11551212;
 RA Elssner T., Engemann C., Baumgart K., Kleber H.-P.;
 RT "Involvement of coenzyme A esters and two new enzymes, an enoyl-CoA
 RT hydratase and a CoA-transferase, in the hydration of crotonobetaine
 RT to L-carnitine by Escherichia coli.";
 RL Biochemistry 40:11140-11148(2001).
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitiny-CoA
 CC to crotonobetainyl-CoA.
 CC -!- CATALYTIC ACTIVITY: L-carnitiny-CoA = H(2)O + crotonobetainyl-
 CC CoA.
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to
 CC gamma-butyrobetaine); second step.
 CC -!- SIMILARITY: BELONGS TO THE ENOYL-COA HYDRATASE/ISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D10483; BAB96605.1; ALT_INIT.
 DR EMBL; X73904; CAA52114.1; ALT_INIT.
 DR EMBL; AE000114; AAC73147.1; ALT_INIT.
 DR HSSP; P14604; 2DUB.
 DR EcoGene; EG11557; caid.
 DR HAMAP; MF_01051; -; 1.
 DR InterPro; IPR001753; EnCoA_hydrase.
 DR Pfam; PF00378; ECH; 1.
 DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
 KW Lyase; Complete proteome.
 FT INIT_MET 0 0
 FT ACT_SITE 110 110 BASE (BY SIMILARITY).
 FT ACT_SITE 130 130 ACID (BY SIMILARITY).
 FT VARIANT 246 246 P -> L (IN STRAIN O44:K74).
 FT CONFLICT 108 108 G -> A (IN REF. 1).
 SQ SEQUENCE 260 AA; 28059 MW; 0DC5E6C3C928A5CC CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;
 Best Local Similarity 72.7%; Pred. No. 7.6;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13
 | | | | |
 Db 59 FSAGWDLKAAA 69

RESULT 6
 CAID_SALTI
 ID CAID_SALTI STANDARD; PRT; 260 AA.
 AC Q8Z9L5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Carnitiny-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA
 DE hydratase).
 GN CAID OR STY0080 OR T0071.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Hague A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitiny-CoA
 CC to crotonobetainyl-CoA (By similarity).
 CC -!- CATALYTIC ACTIVITY: L-carnitiny-CoA = H(2)O + crotonobetainyl-
 CC CoA.
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to
 CC gamma-butyrobetaine); second step.
 CC -!- SIMILARITY: BELONGS TO THE ENOYL-COA HYDRATASE/ISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL627265; CAD01224.1; -.
 DR EMBL; AE016834; AAO67804.1; -.
 DR HAMAP; MF_01051; -; 1.
 DR InterPro; IPR001753; EnCoA_hydratase.
 DR Pfam; PF00378; ECH; 1.
 DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
 KW Lyase; Complete proteome.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT ACT_SITE 110 110 BASE (BY SIMILARITY).

FT ACT_SITE 130 130 ACID (BY SIMILARITY).
SQ SEQUENCE 260 AA; 27945 MW; CE38ED88106A6AA5 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;
Best Local Similarity 72.7%; Pred. No. 7.6;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13
| | | | |
Db 59 FSAGWDLKAAA 69

RESULT 7

CAID_SALTY

ID CAID_SALTY STANDARD; PRT; 260 AA.
AC Q8ZRX5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE CarnitinyL-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA
DE hydratase).
GN CAID OR STM0070.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitinyL-CoA
CC to crotonobetainyl-CoA (By similarity).
CC -!- CATALYTIC ACTIVITY: L-carnitinyL-CoA = H(2)O + crotonobetainyl-
CC CoA.
CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to
CC gamma-butyrobetaine); second step.
CC -!- SIMILARITY: BELONGS TO THE ENOYL-COA HYDRATASE/ISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008696; AAL19034.1; -.
DR StyGene; SG?????; caid.
DR HAMAP; MF_01051; -; 1.

DR InterPro; IPR001753; EnCoA_hydratase.
 DR Pfam; PF00378; ECH; 1.
 DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
 KW Lyase; Complete proteome.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT ACT_SITE 110 110 BASE (BY SIMILARITY).
 FT ACT_SITE 130 130 ACID (BY SIMILARITY).
 SQ SEQUENCE 260 AA; 27975 MW; 30C90296FF6A7BAF CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;
 Best Local Similarity 72.7%; Pred. No. 7.6;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FVAAWTLKAAA 13
 | | | | |
 Db 59 FSAGWDLKAAA 69

RESULT 8

CAID_SHIFL

ID CAID_SHIFL STANDARD; PRT; 260 AA.
 AC P59395;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Carnitiny-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA
 DE hydratase).
 GN CAID OR SF0033.
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitiny-CoA
 CC to crotonobetainyl-CoA (By similarity).
 CC -!- CATALYTIC ACTIVITY: L-carnitiny-CoA + H(2)O = crotonobetainyl-
 CC CoA.
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to gamma-
 CC butyrobetaine); second step.
 CC -!- SIMILARITY: BELONGS TO THE ENOYL-COA HYDRATASE/ISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AE015040; AAN41699.1; ALT_INIT.
DR HAMAP; MF_01051; -, 1.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
KW Lyase.
FT INIT_MET 0 0 BY SIMILARITY.
FT ACT_SITE 110 110 BASE (BY SIMILARITY).
FT ACT_SITE 130 130 ACID (BY SIMILARITY).
SQ SEQUENCE 260 AA; 28089 MW; 0DC5FC696839A5CC CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;
Best Local Similarity 72.7%; Pred. No. 7.6;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FVAAWTLKAAA 13
| | | | |
Db 59 FSAGWDLKAAA 69

RESULT 9

LACB_MACGI

ID LACB_MACGI STANDARD; PRT; 155 AA.
AC P11944;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Beta-lactoglobulin.
GN LGB.
OS Macropus giganteus (Eastern gray kangaroo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9317;
RN [1]
RP SEQUENCE.
RX MEDLINE=87299024; PubMed=3620116;
RA Godovac-Zimmermann J., Shaw D.;
RT "Beta-lactoglobulin identified in marsupial milk. The primary
RT structure, binding site and possible function of beta-lactoglobulin
RT from eastern grey kangaroo (Macropus giganteus).";
RL Biol. Chem. Hoppe-Seyler 368:879-886(1987).
CC -!- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT
CC BINDS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF
CC THAT MOLECULE.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to the lipocalin family.
DR PIR; A29699; A29699.
DR HSSP; P02754; 1BSQ.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR Pfam; PF00061; lipocalin; 1.
DR PROSITE; PS00213; LIPOCALIN; FALSE_NEG.
KW Milk; Whey; Retinol-binding; Transport; Lipocalin.
FT DISULFID 61 153 BY SIMILARITY.
FT DISULFID 103 115 BY SIMILARITY.
SQ SEQUENCE 155 AA; 18340 MW; 7BD7EA4A191530E5 CRC64;

Query Match 59.4%; Score 38; DB 1; Length 155;
Best Local Similarity 58.3%; Pred. No. 7;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13
||| :| |: ||
Db 14 KFGVSWYLREAA 25

RESULT 10

LACB_MACEU

ID LACB_MACEU STANDARD; PRT; 174 AA.

AC Q29614;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Beta-lactoglobulin precursor.

GN LGB OR BLG.

OS Macropus eugenii (Tammar wallaby).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.

OX NCBI_TaxID=9315;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91197370; PubMed=1707629;

RA Collet C., Joseph R., Nicholas K.R.;

RT "A marsupial beta-lactoglobulin gene: characterization and prolactin-
RT dependent expression.";

RL J. Mol. Endocrinol. 6:9-16(1991).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=95314595; PubMed=7794241;

RA Collet C., Joseph R.;

RT "Exon organization and sequence of the genes encoding alpha-
RT lactalbumin and beta-lactoglobulin from the tammar wallaby
RT (Macropodidae, Marsupialia).";

RL Biochem. Genet. 33:61-72(1995).

RN [3]

RP SEQUENCE OF 47-174 FROM N.A.

RC TISSUE=Mammary gland;

RX MEDLINE=95085593; PubMed=7993373;

RA Collet C., Joseph R.;

RT "The identification of nuclear and mitochondrial genes by sequencing
RT randomly chosen clones from a marsupial mammary gland cDNA library.";

RL Biochem. Genet. 32:181-190(1994).

CC -!- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT
CC BINDS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF
CC THAT MOLECULE.

CC -!- SUBUNIT: Monomer (By similarity).

CC -!- SIMILARITY: Belongs to the lipocalin family.

CC

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; L14959; AAA31597.1; -.

DR EMBL; L14954; AAA31597.1; JOINED.

DR EMBL; L14955; AAA31597.1; JOINED.

DR EMBL; L14956; AAA31597.1; JOINED.

DR EMBL; L14957; AAA31597.1; JOINED.

DR EMBL; L14958; AAA31597.1; JOINED.

DR EMBL; X15212; CAA33282.1; -.

DR HSSP; P02754; 1BSY.

DR InterPro; IPR000566; Lipocln_cytFABP.

DR Pfam; PF00061; lipocalin; 1.

DR PROSITE; PS00213; LIPOCALIN; FALSE_NEG.

KW Milk; Whey; Retinol-binding; Transport; Lipocalin; Signal.

FT SIGNAL 1 18 BY SIMILARITY.

FT CHAIN 19 174 BETA-LACTOGLOBULIN.

FT DISULFID 79 172 BY SIMILARITY.

FT DISULFID 122 134 BY SIMILARITY.

FT CONFLICT 76 76 A -> T (IN REF. 3).

SQ SEQUENCE 174 AA; 20219 MW; 6C5901051F99C991 CRC64;

Query Match 59.4%; Score 38; DB 1; Length 174;

Best Local Similarity 58.3%; Pred. No. 7.9;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13

||| :| |: ||

Db 32 KPVGSWYLREAA 43

RESULT 11

UBCX_PICPA

ID UBCX_PICPA STANDARD; PRT; 204 AA.

AC P49428;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Ubiquitin-conjugating enzyme E2-24 kDa (EC 6.3.2.19) (Ubiquitin-protein ligase) (Ubiquitin carrier protein) (Peroxin-4).

GN PEX4 OR PAS4.

OS Pichia pastoris (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Pichia.

OX NCBI_TaxID=4922;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94342381; PubMed=8063827;

RA Crane D.I., Kalish J.E., Gould S.J.;

RT "The Pichia pastoris PAS4 gene encodes a ubiquitin-conjugating enzyme required for peroxisome assembly.";

RL J. Biol. Chem. 269:21835-21844(1994).

CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER PROTEINS. ESSENTIAL FOR PEROXISOME BIOGENESIS.

CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP + diphosphate + protein N-ubiquityllysine.

CC -!- PATHWAY: Ubiquitin conjugation; second step.

CC -!- SUBCELLULAR LOCATION: Peroxisomal.

CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
 CC thiolester formation.
 CC -!- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.
 CC STRONGEST, TO YEAST UBC10.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U12511; AAA53634.1; -.
 DR PIR; A53848; A53848.
 DR HSSP; Q95044; 2E2C.
 DR InterPro; IPR000608; UBQ_conjugat.
 DR Pfam; PF00179; UQ_con; 1.
 DR ProDom; PD000461; UBQ_conjugat; 1.
 DR SMART; SM00212; UBCC; 1.
 DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ubl conjugation pathway; Ligase; Multigene family; Peroxisome.
 FT BINDING 133 133 UBIQUITIN (BY SIMILARITY).
 FT MUTAGEN 133 133 C->S,A: LOSS OF ACTIVITY.
 SQ SEQUENCE 204 AA; 23565 MW; 46680ABD73121F6C CRC64;

Query Match 59.4%; Score 38; DB 1; Length 204;
 Best Local Similarity 58.3%; Pred. No. 9.1;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAA 12
 ||: |||| :|
 Db 139 AKWTPAWTLSSA 150

RESULT 12

VGLM_HSVBC

ID VGLM_HSVBC STANDARD; PRT; 411 AA.
 AC P52370;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Glycoprotein M.
 GN GM OR UL10.
 OS Bovine herpesvirus type 1 (strain Cooper).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95313343; PubMed=7793062;
 RA Vlcek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
 RA Letchworth G.J., Schwyzer M.;
 RT "Nucleotide sequence analysis of a 30-kb region of the bovine
 RT herpesvirus 1 genome which exhibits a colinear gene arrangement with
 RT the UL21 to UL4 genes of herpes simplex virus.";

RL Virology 210:100-108(1995).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN M.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z48053; CAA88123.1; -.
 DR PIR; S61245; S61245.
 DR InterPro; IPR000785; Herpes_glycop.
 DR Pfam; PF01528; Herpes_glycop; 1.
 DR PRINTS; PR00333; HSVINTEGRALMP.
 KW Transmembrane; Glycoprotein.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 89 109 POTENTIAL.
 FT TRANSMEM 133 153 POTENTIAL.
 FT TRANSMEM 159 179 POTENTIAL.
 FT TRANSMEM 213 233 POTENTIAL.
 FT TRANSMEM 249 269 POTENTIAL.
 FT TRANSMEM 277 297 POTENTIAL.
 FT TRANSMEM 318 338 POTENTIAL.
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 411 AA; 43029 MW; 20F156DA9F40158C CRC64;

Query Match 59.4%; Score 38; DB 1; Length 411;
 Best Local Similarity 77.8%; Pred. No. 18;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VAAWTLKAA 12
 :|||||:|
 Db 141 LAAWTLQAA 149

RESULT 13

APGM_PYRAB

ID APMG_PYRAB STANDARD; PRT; 410 AA.
 AC Q9V2M6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE 2,3-bisphosphoglycerate-independent phosphoglycerate mutase
 DE (EC 5.4.2.1) (Phosphoglyceromutase) (BPG-independent PGAM) (aPGAM).
 GN APMG OR PYRAB00490 OR PAB2318.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GE5 / Orsay;
 RX PubMed=12622808;
 RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,

RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
 RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
 RT "An integrated analysis of the genome of the hyperthermophilic
 RT archaeon *Pyrococcus abyssi*.";
 RL Mol. Microbiol. 47:1495-1512(2003).
 CC -!- FUNCTION: Catalyzes the interconversion of 2-phosphoglycerate and
 CC 3-phosphoglycerate (By similarity).
 CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = 3-phospho-D-glycerate.
 CC -!- SIMILARITY: BELONGS TO THE BPG-INDEPENDENT PHOSPHOGLYCERATE MUTASE
 CC FAMILY. A-PGAM SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AJ248283; CAB48972.1; -.
 DR PIR; E75190; E75190.
 DR HAMAP; MF_01402; -, 1.
 DR InterPro; IPR004456; BcpB.
 DR InterPro; IPR006124; Metalloenzyme.
 DR Pfam; PF01676; Metalloenzyme; 1.
 DR ProDom; PD004704; BcpB; 1.
 DR TIGRFAMs; TIGR00306; bcpB; 1.
 KW Isomerase; Complete proteome.
 SQ SEQUENCE 410 AA; 45201 MW; 91C8A89637608826 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 410;
 Best Local Similarity 58.3%; Pred. No. 26;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 KFVAAWTLKAAA 13
 || | :|||
 Db 235 KFTEQWKVKAAA 246

RESULT 14

APGM_PYRFU

ID APMG_PYRFU STANDARD; PRT; 411 AA.
 AC P58814;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 2,3-bisphosphoglycerate-independent phosphoglycerate mutase
 DE (EC 5.4.2.1) (Phosphoglyceromutase) (BPG-independent PGAM) (aPGAM).
 GN APMG OR PF1959.
 OS *Pyrococcus furiosus*.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC *Pyrococcus*.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;

RT "The complete sequence of the *Pyrococcus furiosus* genome.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP CHARACTERIZATION.
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
 RX MEDLINE=22071963; PubMed=12076796;
 RA van der Oost J., Huynen M.A., Verhees C.H.;
 RT "Molecular characterization of phosphoglycerate mutase in archaea.";
 RL FEMS Microbiol. Lett. 212:111-120(2002).
 CC -!- FUNCTION: Catalyzes the interconversion of 2-phosphoglycerate and
 CC 3-phosphoglycerate.
 CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = 3-phospho-D-glycerate.
 CC -!- COFACTOR: Magnesium (Probable).
 CC -!- ENZYME REGULATION: Inhibited to approximately 20% by EDTA.
 CC -!- SUBUNIT: Homotetramer (Probable).
 CC -!- MISCELLANEOUS: Optimal pH is 8.0.
 CC -!- SIMILARITY: BELONGS TO THE BPG-INDEPENDENT PHOSPHOGLYCERATE MUTASE
 CC FAMILY. A-PGAM SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE010289; AAL82083.1; -.
 DR HAMAP; MF_01402; -; 1.
 DR InterPro; IPR004456; BcpB.
 DR InterPro; IPR006124; Metalloenzyme.
 DR Pfam; PF01676; Metalloenzyme; 1.
 DR ProDom; PD004704; BcpB; 1.
 DR TIGRFAMS; TIGR00306; bcpB; 1.
 KW Isomerase; Magnesium; Complete proteome.
 SQ SEQUENCE 411 AA; 45314 MW; 870587E630C2B104 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 411;
 Best Local Similarity 58.3%; Pred. No. 27;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 KFVAAWTLKAAA 13
 || | :|||
 Db 236 KFTEQWKVKAAA 247

RESULT 15

YAHA_ECOLI

ID YAHA_ECOLI STANDARD; PRT; 362 AA.
 AC P21514; P75689;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein yahA.
 GN YAHA OR B0315.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
 RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
 RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE OF 1-126 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=92065800; PubMed=1956285;
 RA Lamark T., Kaasen E., Eshoo M.W., Falkenberg P., McDougall J.,
 RA Strom A.R.;
 RT "DNA sequence and analysis of the bet genes encoding the
 RT osmoregulatory choline-glycine betaine pathway of Escherichia coli.";
 RL Mol. Microbiol. 5:1049-1064(1991).
 CC -!- SIMILARITY: Contains 1 EAL domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000138; AAC73418.1; -.
 DR EMBL; U73857; AAB18041.1; ALT_INIT.
 DR EMBL; X52905; CAA37089.1; -.
 DR PIR; C64758; C64758.
 DR EcoGene; EG11236; yahA.
 DR InterPro; IPR001633; EAL.
 DR InterPro; IPR000792; HTH_LuxR.
 DR Pfam; PF00563; EAL; 1.
 DR Pfam; PF00196; GerE; 1.
 DR PRINTS; PR00038; HTHLUXR.
 DR SMART; SM00052; DUF2; 1.
 DR SMART; SM00421; HTH_LUXR; 1.
 DR PROSITE; PS50883; EAL; 1.
 KW Hypothetical protein; Complete proteome.
 FT DOMAIN 106 360 EAL.
 SQ SEQUENCE 362 AA; 40725 MW; 4BCA091840F27A0D CRC64;

Query Match 56.2%; Score 36; DB 1; Length 362;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFVAAWTLKA 11
 ||:: | :||
Db 351 KFISEWVMKA 360

Search completed: December 13, 2003, 13:35:08
Job time : 1.44444 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 0.809524 Seconds
(without alignments)
1544.356 Million cell updates/sec

Title: US-09-785-215-19
Perfect score: 64
Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	60.9	146	2	D70766	hypothetical prote
2	39	60.9	261	2	AH0510	carnitine racemase
3	39	60.9	297	1	D64724	carnitine racemase
4	39	60.9	297	2	G85484	carnitine racemase
5	39	60.9	297	2	G90633	carnitine racemase
6	38	59.4	155	2	A29699	beta-lactoglobulin
7	38	59.4	204	2	A53848	ubiquitin-conjugat
8	38	59.4	352	2	AE3610	probable transmembr
9	38	59.4	411	2	S61245	probable virion gl
10	38	59.4	639	2	D95884	hypothetical prote
11	38	59.4	758	2	B82122	ferrous iron trans
12	37	57.8	309	2	F70796	hypothetical prote
13	37	57.8	368	2	G87458	conserved hypothet

14	37	57.8	410	2	E75190	probable phosphono
15	37	57.8	712	2	AG1888	Na ⁺ /H ⁺ antiporter
16	36	56.2	156	2	T02806	probable membrane
17	36	56.2	317	2	F87336	conserved hypothet
18	36	56.2	330	2	G89771	lipoprotein [impor
19	36	56.2	357	2	T12379	NADH2 dehydrogenas
20	36	56.2	362	2	C90674	hypothetical prote
21	36	56.2	362	2	C64758	yahA protein - Esc
22	36	56.2	363	2	B82532	GTP-binding protei
23	36	56.2	365	2	F85524	hypothetical prote
24	36	56.2	412	2	T22076	hypothetical prote
25	36	56.2	569	2	T00851	hypothetical prote
26	36	56.2	690	2	T23399	hypothetical prote
27	36	56.2	776	1	JH0570	transferrin recept
28	36	56.2	1031	2	D88912	protein T06A10.1 [
29	36	56.2	1031	2	T33655	hypothetical prote
30	36	56.2	2076	2	T28915	hypothetical prote
31	35	54.7	120	2	AF0299	probable membrane
32	35	54.7	128	2	C48552	hypothetical prote
33	35	54.7	226	2	G81894	probable integral
34	35	54.7	226	2	D81125	hypothetical prote
35	35	54.7	301	2	B90027	hypothetical prote
36	35	54.7	388	2	AF2718	ribonuclease D [im
37	35	54.7	429	2	B97500	ribonuclease D (PA
38	35	54.7	550	2	A57519	parathyroid hormon
39	35	54.7	587	2	T01829	hypothetical prote
40	35	54.7	639	2	S03547	hypothetical prote
41	35	54.7	659	2	A72729	hypothetical prote
42	35	54.7	723	2	T14605	probable cell divi
43	35	54.7	730	2	T44246	translation elonga
44	35	54.7	750	2	A56881	prostate-specific
45	35	54.7	773	2	A82998	probable oxidoredu

ALIGNMENTS

RESULT 1

D70766

hypothetical protein Rv2081c - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C;Accession: D70766

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaiia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: D70766

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA
A;Residues: 1-146 <COL>
A;Cross-references: GB:Z73966; GB:AL123456; NID:g3261577; PIDN:CAA98193.1;
PID:e247091; PID:g1370247
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv2081c

Query Match 60.9%; Score 39; DB 2; Length 146;
Best Local Similarity 63.6%; Pred. No. 8.2;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13
||| || :||:
Db 10 FVAIWTTARAAS 20

RESULT 2

AH0510

carnitine racemase [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhi (strain CT18)

C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi

A;Note: this species has also been called *Salmonella typhi*

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C;Accession: AH0510

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; Mungall, K.L.; Bentley, S.D.; Holden, M.T.G.; Sebaihia, M.; Baker, S.; Basham, D.; Brooks, K.; Chillingworth, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, J.; Feltwell, T.; Hamlin, N.; Haque, A.; Hien, T.T.; Holroyd, S.; Jagels, K.; Krogh, A.; Larsen, T.S.; Leather, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.G.

A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18.

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AH0510

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-261 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD01224.1; PID:g16501353; GSPDB:GN00176

C;Genetics:

A;Gene: STY0080

C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology

Query Match 60.9%; Score 39; DB 2; Length 261;
Best Local Similarity 72.7%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13
| | | |||||
Db 60 FSAGWDLKAAA 70

RESULT 3

D64724

carnitine racemase (EC 5.-.-.-) - Escherichia coli (strain K-12)
 C;Species: Escherichia coli
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
 C;Accession: D64724; S40557; I41014
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.;
 Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor,
 J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503
 A;Accession: D64724
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-297 <BLAT>
 A;Cross-references: GB:AE000114; GB:U00096; NID:g1786217; PIDN:AAC73147.1;
 PID:g1786220; UWGP:b0036
 A;Experimental source: strain K-12, substrain MG1655
 R;Yura, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono,
 K.; Mizobuchi, K.; Nakata, A.
 submitted to the EMBL Data Library, December 1992
 A;Description: Systematic sequencing of the Escherichia coli genome: analysis of
 the 0-2.4min region.
 A;Reference number: S40531
 A;Accession: S40557
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-144, 'A', 146-297 <YUR>
 A;Cross-references: EMBL:D10483; NID:g216434; PIDN:BAA01312.1; PID:g216461
 R;Eichler, K.; Bourgis, F.; Buchet, A.; Kleber, H.P.; Mandrand-Berthelot, M.A.
 Mol. Microbiol. 13, 775-786, 1994
 A;Title: Molecular characterization of the cai operon necessary for carnitine
 metabolism in Escherichia coli.
 A;Reference number: I41010; MUID:95115548; PMID:7815937
 A;Accession: I41014
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-2, 'R', 4-13, 'A', 15-282, 'L', 284-297 <RES>
 A;Cross-references: EMBL:X73904; NID:g563860; PIDN:CAA52114.1; PID:g563865
 A;Experimental source: strain 044 K74
 C;Genetics:
 A;Gene: caiD
 C;Function:
 A;Pathway: carnitine metabolism
 C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology
 C;Keywords: isomerase
 F;60-212/Domain: enoyl-CoA hydratase homology <ECH>

Query Match 60.9%; Score 39; DB 1; Length 297;
 Best Local Similarity 72.7%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13
 | | | | |
 Db 96 FSAGWDLKAAA 106

G85484

carnitine racemase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002

C;Accession: G85484

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Posfai, G.; Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, J.; Anantharaman, T.S.; Lin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.
Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: G85484

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-297 <STO>

A;Cross-references: GB:AE005174; NID:g12512723; PIDN:AAG54339.1; GSPDB:GN00145; UWGP:Z0042

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: caiD

C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology

Query Match 60.9%; Score 39; DB 2; Length 297;
Best Local Similarity 72.7%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13
| | | | |
Db 96 FSAGWDLKAAA 106

RESULT 5

G90633

carnitine racemase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002

C;Accession: G90633

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Ohtsubo, E.; Nakayama, K.; Murata, T.; Tanaka, M.; Tobe, T.; Iida, T.; Takami, H.; Honda, T.; Sasakawa, C.; Ogasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: G90633

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-297 <HAY>

A;Cross-references: GB:BA000007; PIDN:BAB33462.1; PID:g13359495; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECs0039

C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology

Query Match 60.9%; Score 39; DB 2; Length 297;
Best Local Similarity 72.7%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13
| | | | |
Db 96 FSAGWDLKAAA 106

RESULT 6

A29699

beta-lactoglobulin - eastern gray kangaroo

C;Species: Macropus giganteus (eastern gray kangaroo)

C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 18-Jul-1997

C;Accession: A29699

R;Godovac-Zimmermann, J.; Shaw, D.

Biol. Chem. Hoppe-Seyler 368, 879-886, 1987

A;Title: The primary structure, binding site and possible function of beta-lactoglobulin from eastern grey kangaroo (Macropus giganteus).

A;Reference number: A29699; MUID:87299024; PMID:3620116

A;Accession: A29699

A;Molecule type: protein

A;Residues: 1-155 <GOD>

C;Superfamily: lipocalin; lipocalin homology

F;10-155/Domain: lipocalin homology <LIP>

Query Match 59.4%; Score 38; DB 2; Length 155;
Best Local Similarity 58.3%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13
| | | : | | : | |
Db 14 KFGVSWYLREAA 25

RESULT 7

A53848

ubiquitin-conjugating enzyme - yeast (Pichia pastoris)

C;Species: Pichia pastoris

C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 17-Mar-2000

C;Accession: A53848

R;Crane, D.I.; Kalish, J.E.; Gould, S.J.

J. Biol. Chem. 269, 21835-21844, 1994

A;Title: The Pichia pastoris PAS4 gene encodes a ubiquitin-conjugating enzyme required for peroxisome assembly.

A;Reference number: A53848; MUID:94342381; PMID:8063827

A;Accession: A53848

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-204 <CRA>

A;Cross-references: GB:U12511; NID:g531386; PIDN:AAA53634.1; PID:g531387

C;Genetics:

A;Gene: PAS4

C;Superfamily: yeast ubiquitin-protein ligase UBC1

Query Match 59.4%; Score 38; DB 2; Length 204;
Best Local Similarity 58.3%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAA 12
||: |||| :|
Db 139 AKWTPAWTLSSA 150

RESULT 8

AE3610

probable transmembrane protein [imported] - *Brucella melitensis* (strain 16M)

C;Species: *Brucella melitensis*

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C;Accession: AE3610

R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.;

Ivanova, N.; Anderson, I.; Bhattacharyya, A.; Lykidis, A.; Reznik, G.;

Jablonski, L.; Larsen, N.; D'Souza, M.; Bernal, A.; Mazur, M.; Goltsman, E.;

Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesson, J.J.; Haselkorn,

R.; Kyrpides, N.; Overbeek, R.

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*.

A;Reference number: AD3252; PMID:11756688

A;Accession: AE3610

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-352 <KUR>

A;Cross-references: GB:AE008918; PIDN:AAL54048.1; PID:g17985002; GSPDB:GN00191

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BMEII0806

A;Map position: II

Query Match 59.4%; Score 38; DB 2; Length 352;
Best Local Similarity 63.6%; Pred. No. 29;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKA 11
| |:||||:| |
Db 331 AMFIAAWSLLA 341

RESULT 9

S61245

probable virion glycoprotein M (gM) - bovine herpesvirus 1

C;Species: bovine herpesvirus 1

C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 26-Aug-1999

C;Accession: S61245

R;Vlcek, C.; Benes, V.; Lu, Z.; Kutish, G.F.; Paces, V.; Rock, D.; Letchworth,

G.J.; Schwyzer, M.

submitted to the EMBL Data Library, January 1995

A;Description: Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus 1 genome which exhibits a colinear gene arrangement with the UL21 to UL4 genes of herpes simplex virus.

A;Reference number: S61233

A;Accession: S61245

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-411 <VLC>
A;Cross-references: EMBL:Z48053; NID:g971311; PIDN:CAA88123.1; PID:g971324
C;Superfamily: herpesvirus 51K protein

Query Match 59.4%; Score 38; DB 2; Length 411;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VAAWTLKAA 12
:|||||:
Db 141 LAAWTLQAA 149

RESULT 10

D95884

hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021)
megaplasmid pSymB

C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C;Accession: D95884

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernandez-Lucas, I.; Becker, A.; Cowie, A.; Gouzy, J.; Golding, B.; Puhler, A.

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endosymbiont Sinorhizobium meliloti.

A;Reference number: A95842; MUID:21396508; PMID:11481431

A;Accession: D95884

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-639 <KUR>

A;Cross-references: GB:AL591985; PIDN:CAC48740.1; PID:g15140213; GSPDB:GN00167

A;Experimental source: strain 1021, megaplasmid pSymB

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, F.; Barnett, M.J.; Becker, A.; Boistard, P.; Bothe, G.; Boutry, M.; Bowser, L.; Buhrmester, J.; Cadieu, E.; Capela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; Gloux, S.; Godrie, T.; Goffeau, A.; Golding, B.; Gouzy, J.; Gurjal, M.; Hernandez-Lucas, I.; Hong, A.; Huizar, L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, V.; Masuy, D.; Palm, C.; Peck, M.C.; Pohl, T.M.; Portetelle, D.; Purnelle, B.; Ramsperger, U.; Surzycki, R.; Thebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.; Batut, J.

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: SMb20354

A;Genome: plasmid

Query Match 59.4%; Score 38; DB 2; Length 639;
Best Local Similarity 69.2%; Pred. No. 52;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
 | | | | | | |
 Db 491 ALFGAAWTLAYAA 503

RESULT 11

B82122

ferrous iron transport protein B VC2077 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)

C;Species: *Vibrio cholerae*

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 24-Aug-2001

C;Accession: B82122

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill, S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.; Utterback, T.; Fleishmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.; Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M. Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: B82122

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-758 <HEI>

A;Cross-references: GB:AE004281; GB:AE003852; NID:g9656616; PIDN:AAF95223.1; GSPDB:GN00126; TIGR:VC2077

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC2077

A;Map position: 1

C;Superfamily: ferrous iron transport protein B; translation elongation factor Tu homology

Query Match 59.4%; Score 38; DB 2; Length 758;
 Best Local Similarity 50.0%; Pred. No. 61;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAA 12
 | : | : | : |
 Db 694 ARFIAVWTMGLA 705

RESULT 12

F70796

hypothetical protein Rv3725 - *Mycobacterium tuberculosis* (strain H37RV)

C;Species: *Mycobacterium tuberculosis*

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C;Accession: F70796

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaiia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: F70796

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-309 <COL>

A;Cross-references: GB:AL022121; GB:AL123456; NID:g3261559; PIDN:CAA18047.1; PID:g2960149

A;Experimental source: strain H37Rv

C;Genetics:

A;Gene: Rv3725

C;Superfamily: dihydrokaempferol 4-reductase

Query Match 57.8%; Score 37; DB 2; Length 309;
Best Local Similarity 63.6%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13
|| || || |
Db 16 FVGGWTAIA 26

RESULT 13

G87458

conserved hypothetical protein CC1690 [imported] - *Caulobacter crescentus*

C;Species: *Caulobacter crescentus*

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C;Accession: G87458

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.F.; Alley, M.; Ohta, N.; Maddock, J.R.; Potocka, I.; Nelson, W.C.; Newton, A.; Stephens, C.; Phadke, N.D.; Ely, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonay, J.F.; Smit, J.; Craven, M.; Khouri, H.; Shetty, J.; Berry, K.; Utterback, T.; Tran, K.; Wolf, A.; Vamathevan, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of *Caulobacter crescentus*.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: G87458

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-368 <STO>

A;Cross-references: GB:AE005673; NID:g13423101; PIDN:AAK23667.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC1690

Query Match 57.8%; Score 37; DB 2; Length 368;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AKFVAAWT 8
| |||||
Db 342 APFVAAWT 349

RESULT 14

E75190

probable phosphonopyruvate decarboxylase PAB2318 - *Pyrococcus abyssi* (strain Orsay)

C;Species: *Pyrococcus abyssi*

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C;Accession: E75190

R;anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A;Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure and evolution.

A;Reference number: A75001

A;Accession: E75190

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-410 <KAW>

A;Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB48972.1; PID:g5457481

A;Experimental source: strain Orsay

C;Genetics:

A;Gene: PAB2318

C;Superfamily: phosphonopyruvate decarboxylase

Query Match 57.8%; Score 37; DB 2; Length 410;
Best Local Similarity 58.3%; Pred. No. 51;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13
|| | :|||
Db 235 KFTEQWKVKAAA 246

RESULT 15

AG1888

Na⁺/H⁺ antiporter [imported] - *Nostoc* sp. (strain PCC 7120)

C;Species: *Nostoc* sp. PCC 7120

A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C;Accession: AG1888

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120.

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AG1888

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-712 <KUR>

A;Cross-references: GB:BA000019; PIDN:BAB72614.1; PID:g17130002; GSPDB:GN00179

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: alr0656

Query Match 57.8%; Score 37; DB 2; Length 712;
Best Local Similarity 60.0%; Pred. No. 87;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLK 10
:|:|:|
Db 323 SKFIAAWLAK 332

Search completed: December 13, 2003, 13:30:57
Job time : 3.80952 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:30:00 ; Search time 1.47619 Seconds
(without alignments)
1637.856 Million cell updates/sec

Title: US-09-785-215-19
Perfect score: 64
Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match Length	DB	ID		Description

1	64	100.0	13	10	US-09-894-018-69	Sequence 69, Appl
2	64	100.0	13	10	US-09-949-375A-18	Sequence 18, Appl
3	64	100.0	13	10	US-09-785-215-19	Sequence 19, Appl
4	64	100.0	13	12	US-10-223-809A-17	Sequence 17, Appl
5	64	100.0	13	12	US-10-295-074-7	Sequence 7, Appli
6	64	100.0	13	12	US-10-295-074-20	Sequence 20, Appl
7	64	100.0	13	12	US-10-371-069-52	Sequence 52, Appl
8	64	100.0	13	12	US-10-371-645-52	Sequence 52, Appl
9	64	100.0	13	12	US-10-371-260-52	Sequence 52, Appl
10	64	100.0	55	10	US-09-949-375A-12	Sequence 12, Appl
11	64	100.0	80	10	US-09-894-018-97	Sequence 97, Appl
12	64	100.0	80	12	US-10-371-069-22	Sequence 22, Appl
13	64	100.0	80	12	US-10-371-645-22	Sequence 22, Appl
14	64	100.0	80	12	US-10-371-260-22	Sequence 22, Appl
15	64	100.0	98	10	US-09-894-018-101	Sequence 101, App
16	64	100.0	106	10	US-09-894-018-95	Sequence 95, Appl
17	64	100.0	107	10	US-09-894-018-93	Sequence 93, Appl
18	64	100.0	107	10	US-09-894-018-103	Sequence 103, App
19	64	100.0	118	12	US-10-371-069-8	Sequence 8, Appli
20	64	100.0	118	12	US-10-371-645-8	Sequence 8, Appli
21	64	100.0	118	12	US-10-371-260-8	Sequence 8, Appli
22	64	100.0	119	12	US-10-371-069-10	Sequence 10, Appl
23	64	100.0	119	12	US-10-371-645-10	Sequence 10, Appl
24	64	100.0	119	12	US-10-371-260-10	Sequence 10, Appl
25	64	100.0	122	12	US-10-371-069-39	Sequence 39, Appl
26	64	100.0	122	12	US-10-371-645-39	Sequence 39, Appl
27	64	100.0	122	12	US-10-371-260-39	Sequence 39, Appl
28	64	100.0	123	10	US-09-894-018-109	Sequence 109, App
29	64	100.0	130	10	US-09-894-018-99	Sequence 99, Appl
30	64	100.0	132	12	US-10-371-069-37	Sequence 37, Appl
31	64	100.0	132	12	US-10-371-645-37	Sequence 37, Appl
32	64	100.0	132	12	US-10-371-260-37	Sequence 37, Appl
33	64	100.0	136	12	US-10-371-069-12	Sequence 12, Appl
34	64	100.0	136	12	US-10-371-645-12	Sequence 12, Appl
35	64	100.0	136	12	US-10-371-260-12	Sequence 12, Appl
36	64	100.0	138	12	US-10-371-069-4	Sequence 4, Appli
37	64	100.0	138	12	US-10-371-645-4	Sequence 4, Appli
38	64	100.0	138	12	US-10-371-260-4	Sequence 4, Appli
39	64	100.0	144	10	US-09-894-018-129	Sequence 129, App
40	64	100.0	147	10	US-09-894-018-131	Sequence 131, App
41	64	100.0	148	10	US-09-894-018-127	Sequence 127, App
42	64	100.0	152	12	US-10-371-069-20	Sequence 20, Appl
43	64	100.0	152	12	US-10-371-645-20	Sequence 20, Appl
44	64	100.0	152	12	US-10-371-260-20	Sequence 20, Appl
45	64	100.0	157	10	US-09-894-018-117	Sequence 117, App

ALIGNMENTS

RESULT 1
 US-09-894-018-69
 ; Sequence 69, Application US/09894018
 ; Patent No. US20020119127A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EPIMMUNE, Inc.

```
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE sequence
US-09-894-018-69
```

```
Query Match          100.0%; Score 64; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      1 AKFVAAWTLKAAA 13
```

RESULT 2

```
US-09-949-375A-18
; Sequence 18, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 17.
US-09-949-375A-18
```

```
Query Match          100.0%; Score 64; DB 10; Length 13;
```

Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
|||
Db 1 AKFVAAWTLKAAA 13

RESULT 3

US-09-785-215-19

; Sequence 19, Application US/09785215
; Publication No. US20020187157A1
; GENERAL INFORMATION:
; APPLICANT: JENSEN, Martin Roland et al.
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID
; FILE REFERENCE: 3631-0107P
; CURRENT APPLICATION NUMBER: US/09/785,215
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial T-cell epitope capable of binding to a large
portion
; OTHER INFORMATION: of MHC Class II molecules in a variety of animals
US-09-785-215-19

Query Match 100.0%; Score 64; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
|||
Db 1 AKFVAAWTLKAAA 13

RESULT 4

US-10-223-809A-17

; Sequence 17, Application US/10223809A
; Publication No. US20030157117A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; APPLICANT: Rasmussen, Peter Birk et al.
; TITLE OF INVENTION: No. US20030157117A1el Method for Down-Regulation of
Amyloid
; FILE REFERENCE: 674542-2008
; CURRENT APPLICATION NUMBER: US/10/223,809A
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: US 60/337,543
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: US 60/373,027
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: DE 2001 01231
; PRIOR FILING DATE: 2001-08-20

; PRIOR APPLICATION NUMBER: DE 2002 0058
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HLA DR binding sequence
US-10-223-809A-17

Query Match 100.0%; Score 64; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
| | | | | | | | | | | | |
Db 1 AKFVAAWTLKAAA 13

RESULT 5

US-10-295-074-7
; Sequence 7, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Pan DR binding peptide (PADRE)
US-10-295-074-7

Query Match 100.0%; Score 64; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
| | | | | | | | | | | | |
Db 1 AKFVAAWTLKAAA 13

RESULT 6

US-10-295-074-20
; Sequence 20, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS

; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Pan DR binding peptide (PADRE)
US-10-295-074-20

Query Match 100.0%; Score 64; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
|||
Db 1 AKFVAAWTLKAAA 13

RESULT 7

US-10-371-069-52
; Sequence 52, Application US/10371069
; Publication No. US20030216342A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.10
; CURRENT APPLICATION NUMBER: US/10/371,069
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE-universal MHC class II epitope
US-10-371-069-52

Query Match 100.0%; Score 64; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
| | | | | | | | | |
Db 1 AKFVAAWTLKAAA 13

RESULT 8

US-10-371-645-52
; Sequence 52, Application US/10371645
; Publication No. US20030216343A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.11
; CURRENT APPLICATION NUMBER: US/10/371,645
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE-universal MHC class II epitope
US-10-371-645-52

Query Match 100.0%; Score 64; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
| | | | | | | | | |
Db 1 AKFVAAWTLKAAA 13

RESULT 9

US-10-371-260-52
; Sequence 52, Application US/10371260
; Publication No. US20030220285A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.

```
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.13
; CURRENT APPLICATION NUMBER: US/10/371,260
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE-universal MHC class II epitope
US-10-371-260-52
```

```
Query Match          100.0%; Score 64; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      1 AKFVAAWTLKAAA 13
```

RESULT 10

```
US-09-949-375A-12
; Sequence 12, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 11.
US-09-949-375A-12
```

```
Query Match          100.0%; Score 64; DB 10; Length 55;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AKFVAAWTLKAAA 13
        |||||
```

Db 13 AKFVAAWTLKAAA 25

RESULT 11

US-09-894-018-97

; Sequence 97, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV.3s2(-3)
US-09-894-018-97

Query Match 100.0%; Score 64; DB 10; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
|||
Db 32 AKFVAAWTLKAAA 44

RESULT 12

US-10-371-069-22

; Sequence 22, Application US/10371069
; Publication No. US20030216342A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.

```

; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.10
; CURRENT APPLICATION NUMBER: US/10/371,069
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SigTh
US-10-371-069-22

```

```

Query Match          100.0%; Score 64; DB 12; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      68 AKFVAAWTLKAAA 80

```

```

RESULT 13
US-10-371-645-22
; Sequence 22, Application US/10371645
; Publication No. US20030216343A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.11
; CURRENT APPLICATION NUMBER: US/10/371,645
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

; FEATURE:
; OTHER INFORMATION: SigTh
US-10-371-645-22

Query Match 100.0%; Score 64; DB 12; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13
|||
Db 68 AKFVAAWTLKAAA 80

RESULT 14

US-10-371-260-22

; Sequence 22, Application US/10371260
; Publication No. US20030220285A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.13
; CURRENT APPLICATION NUMBER: US/10/371,260
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SigTh
US-10-371-260-22

Query Match 100.0%; Score 64; DB 12; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13
|||
Db 68 AKFVAAWTLKAAA 80

RESULT 15

US-09-894-018-101

; Sequence 101, Application US/09894018

```

; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV.PC3
US-09-894-018-101

```

```

Query Match          100.0%; Score 64; DB 10; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      34 AKFVAAWTLKAAA 46

```

```

Search completed: December 13, 2003, 13:36:50
Job time : 1.47619 secs

```

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 0.793651 Seconds
(without alignments)
693.052 Million cell updates/sec

Title: US-09-785-215-19
Perfect score: 64
Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	64	100.0	13	1	US-08-305-871A-23	Sequence 23, Appl
2	64	100.0	13	4	US-08-788-822A-27	Sequence 27, Appl
3	64	100.0	80	4	US-09-311-784A-22	Sequence 22, Appl
4	64	100.0	118	4	US-09-311-784A-8	Sequence 8, Appli
5	64	100.0	119	4	US-09-311-784A-10	Sequence 10, Appl
6	64	100.0	122	4	US-09-311-784A-39	Sequence 39, Appl
7	64	100.0	132	4	US-09-311-784A-37	Sequence 37, Appl
8	64	100.0	136	4	US-09-311-784A-12	Sequence 12, Appl
9	64	100.0	138	4	US-09-311-784A-4	Sequence 4, Appli
10	64	100.0	152	4	US-09-311-784A-20	Sequence 20, Appl
11	64	100.0	164	4	US-09-311-784A-18	Sequence 18, Appl

12	64	100.0	215	4	US-09-311-784A-2	Sequence 2, Appli
13	64	100.0	249	4	US-09-311-784A-16	Sequence 16, Appl
14	64	100.0	253	4	US-09-311-784A-6	Sequence 6, Appli
15	64	100.0	266	4	US-09-311-784A-14	Sequence 14, Appl
16	57	89.1	13	1	US-08-305-871A-22	Sequence 22, Appl
17	57	89.1	13	4	US-08-788-822A-26	Sequence 26, Appl
18	57	89.1	13	4	US-09-692-170C-29	Sequence 29, Appl
19	57	89.1	22	4	US-09-692-170C-17	Sequence 17, Appl
20	57	89.1	22	4	US-09-692-170C-18	Sequence 18, Appl
21	57	89.1	22	4	US-09-692-170C-20	Sequence 20, Appl
22	57	89.1	22	4	US-09-692-170C-21	Sequence 21, Appl
23	57	89.1	22	4	US-09-692-170C-22	Sequence 22, Appl
24	57	89.1	22	4	US-09-692-170C-23	Sequence 23, Appl
25	57	89.1	22	4	US-09-692-170C-24	Sequence 24, Appl
26	57	89.1	23	4	US-09-692-170C-15	Sequence 15, Appl
27	57	89.1	23	4	US-09-692-170C-16	Sequence 16, Appl
28	57	89.1	24	4	US-09-983-019-10	Sequence 10, Appl
29	57	89.1	25	3	US-09-075-257A-17	Sequence 17, Appl
30	57	89.1	25	3	US-09-534-639-17	Sequence 17, Appl
31	57	89.1	25	4	US-09-692-170C-25	Sequence 25, Appl
32	57	89.1	25	4	US-09-692-170C-37	Sequence 37, Appl
33	57	89.1	25	4	US-09-692-170C-44	Sequence 44, Appl
34	57	89.1	27	4	US-09-983-019-4	Sequence 4, Appli
35	57	89.1	27	4	US-09-983-019-7	Sequence 7, Appli
36	57	89.1	28	3	US-09-075-257A-15	Sequence 15, Appl
37	57	89.1	28	3	US-09-075-257A-16	Sequence 16, Appl
38	57	89.1	28	3	US-09-534-639-15	Sequence 15, Appl
39	57	89.1	28	3	US-09-534-639-16	Sequence 16, Appl
40	57	89.1	28	4	US-09-692-170C-35	Sequence 35, Appl
41	57	89.1	28	4	US-09-692-170C-36	Sequence 36, Appl
42	49	76.6	12	4	US-08-788-822A-21	Sequence 21, Appl
43	39	60.9	297	3	US-09-058-489-10	Sequence 10, Appl
44	37	57.8	13	1	US-08-305-871A-20	Sequence 20, Appl
45	37	57.8	13	4	US-08-788-822A-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-08-305-871A-23

; Sequence 23, Application US/08305871A

; Patent No. 5736142

; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro

; APPLICANT: Gaeta, Federico

; APPLICANT: Grey, Howard M.

; APPLICANT: Sidney, John

; APPLICANT: Alexander, Jeffrey L.

; TITLE OF INVENTION: Alteration of Immune Response Using Pan

; TITLE OF INVENTION: DR-Binding Peptides

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California


```

; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/305,871A
; FILING DATE: 14-SEP-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,101
; FILING DATE: 14-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-0062-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..13
; OTHER INFORMATION: /note= "Peptide wherein X is
; OTHER INFORMATION: tyrosine or phenylalanine."
US-08-305-871A-23

```

```

Query Match          100.0%; Score 64; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      1 AKFVAAWTLKAAA 13

```

```

RESULT 2
US-08-788-822A-27
; Sequence 27, Application US/08788822A
; Patent No. 6413935
; GENERAL INFORMATION:
; APPLICANT: Alexander, Jeffrey L.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Induction of Immune Response Against
; TITLE OF INVENTION: Desired Determinants
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP

```

; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/788,822A
 ; FILING DATE: 23-JAN-1997
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/010,510
 ; FILING DATE: 24-JAN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bastian, Kevin L.
 ; REGISTRATION NUMBER: 34,774
 ; REFERENCE/DOCKET NUMBER: 014137-009210US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 27:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 13 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-08-788-822A-27

Query Match 100.0%; Score 64; DB 4; Length 13;
 Best Local Similarity 100.0%; Pred. No. 0.00024;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13
 |||||
 Db 1 AKFVAAWTLKAAA 13

RESULT 3

US-09-311-784A-22

; Sequence 22, Application US/09311784A
 ; Patent No. 6534482
 ; GENERAL INFORMATION:
 ; APPLICANT: Fikes, John D.
 ; APPLICANT: Hermanson, Gary G.
 ; APPLICANT: Sette, Alessandro
 ; APPLICANT: Ishioka, Glenn Y.
 ; APPLICANT: Livingston, Brian
 ; APPLICANT: Chesnut, Robert W.
 ; APPLICANT: Epimmune Inc.
 ; TITLE OF INVENTION: Expression Vectors for Stimulating an
 ; TITLE OF INVENTION: Immune Response and Methods of Using the Same
 ; FILE REFERENCE: 39963-20022.01

```
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SigTh
US-09-311-784A-22
```

```
Query Match          100.0%; Score 64; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      68 AKFVAAWTLKAAA 80
```

RESULT 4

US-09-311-784A-8

```
; Sequence 8, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: KappaLAMP-Th
US-09-311-784A-8
```

```
Query Match          100.0%; Score 64; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AKFVAAWTLKAAA 13
```

Db

68 AKFVAAWTLKAAA 80

RESULT 5

US-09-311-784A-10

; Sequence 10, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: H2M-Th
US-09-311-784A-10

Query Match 100.0%; Score 64; DB 4; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13
|||||||
Db 63 AKFVAAWTLKAAA 75

RESULT 6

US-09-311-784A-39

; Sequence 39, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01

```

; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vector pMIN.1
US-09-311-784A-39

```

```

Query Match          100.0%; Score 64; DB 4; Length 122;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 AKFVAAWTLKAAA 13
        |||||
Db      34 AKFVAAWTLKAAA 46

```

RESULT 7

```

US-09-311-784A-37
; Sequence 37, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vector pMIN.0
US-09-311-784A-37

```

```

Query Match          100.0%; Score 64; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 AKFVAAWTLKAAA 13

```

Db

|||||
34 AKFVAAWTLKAAA 46

RESULT 8

US-09-311-784A-12

; Sequence 12, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: H2O-Th
US-09-311-784A-12

Query Match 100.0%; Score 64; DB 4; Length 136;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
|||||
Db 67 AKFVAAWTLKAAA 79

RESULT 9

US-09-311-784A-4

; Sequence 4, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01

; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: I80T
US-09-311-784A-4

Query Match 100.0%; Score 64; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13
|||||||
Db 126 AKFVAAWTLKAAA 138

RESULT 10

US-09-311-784A-20

; Sequence 20, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ig-betaTh
US-09-311-784A-20

Query Match 100.0%; Score 64; DB 4; Length 152;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13

Db

|||||
71 AKFVAAWTLKAAA 83

RESULT 11

US-09-311-784A-18

; Sequence 18, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ig-alphaTh
US-09-311-784A-18

Query Match 100.0%; Score 64; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13

|||||
Db 74 AKFVAAWTLKAAA 86

RESULT 12

US-09-311-784A-2

; Sequence 2, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01

; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: IiPADRE
US-09-311-784A-2

Query Match 100.0%; Score 64; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13
|||
Db 88 AKFVAAWTLKAAA 100

RESULT 13

US-09-311-784A-16

; Sequence 16, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE-HBV-s
US-09-311-784A-16

Query Match 100.0%; Score 64; DB 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13

Db

|||||
2 AKFVAAWTLKAAA 14

RESULT 14

US-09-311-784A-6

; Sequence 6, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: IiThfull
US-09-311-784A-6

Query Match 100.0%; Score 64; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
|||||
Db 126 AKFVAAWTLKAAA 138

RESULT 15

US-09-311-784A-14

; Sequence 14, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01

; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE-Influenza matrix
US-09-311-784A-14

Query Match 100.0%; Score 64; DB 4; Length 266;
Best Local Similarity 100.0%; Pred. NO. 0.0041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
| | | | | | | | | | | | | |
Db 2 AKFVAAWTLKAAA 14

Search completed: December 13, 2003, 13:31:54
Job time : 0.793651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2003, 13:27:36 ; Search time 2.15873 Seconds
(without alignments)
955.861 Million cell updates/sec

Title: US-09-785-215-19
Perfect score: 64
Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	% Query		Length	DB	ID	Description
	Score	Match				
1	64	100.0	13	16	AAR70250	Pan DR-binding pep
2	64	100.0	13	18	AAW22121	Padre (pan-DR bind
3	64	100.0	13	21	AAB36289	Promiscuous T help
4	64	100.0	13	21	AAV52558	Universal helper T
5	64	100.0	13	22	AAB73644	Pan-DR binding pep
6	64	100.0	13	22	AAB99710	Pan-DR-binding pep
7	64	100.0	13	22	AAB20154	PADRE peptide. Sy
8	64	100.0	13	23	AAE26369	PADRE peptide. Un
9	64	100.0	13	23	ABG31776	Pan DR epitope (PA
10	64	100.0	13	23	AAU80293	Pan DR epitope pep
11	64	100.0	13	24	ABP72696	T cell epitope PAD
12	64	100.0	55	23	AAU80290	Pan DR epitope pep
13	64	100.0	80	21	AAV52548	Murine Ig-kappa si
14	64	100.0	118	21	AAV52541	KappaLAMP-Th fusio
15	64	100.0	119	21	AAV52542	Murine H2-M/multip
16	64	100.0	122	21	AAV52576	Amino acid sequenc
17	64	100.0	132	21	AAV52575	Amino acid sequenc
18	64	100.0	136	21	AAV52543	Murine H2-DO/multi
19	64	100.0	138	21	AAV52539	Murine Ii/multiple
20	64	100.0	152	21	AAV52547	Murine Ig-beta/mul
21	64	100.0	164	21	AAV52546	Murine Ig-alpha/mu
22	64	100.0	215	21	AAV52538	Murine Ii/pan DR e
23	64	100.0	249	21	AAV52545	Pan DR epitope/HBV
24	64	100.0	253	21	AAV52540	Murine Ii/multiple
25	64	100.0	266	21	AAV52544	Pan DR epitope/inf
26	61	95.3	13	22	AAB99711	Pan-DR-binding pep
27	57	89.1	13	16	AAR70249	Pan DR-binding pep
28	57	89.1	13	18	AAW22120	Padre (pan-DR bind
29	57	89.1	13	19	AAW50125	Pan DR binding pep
30	57	89.1	13	19	AAW50126	Pan DR binding pep
31	57	89.1	13	22	AAB99709	Pan-DR-binding pep
32	57	89.1	13	22	AAG84519	Pan-DR-binding epi
33	57	89.1	13	22	AAG88271	Pan-DR-binding epi
34	57	89.1	13	22	AAB46167	PADRE universal T
35	57	89.1	13	22	AAB49066	PADRE T-cell epito
36	57	89.1	13	23	ABJ16183	Zinc transporter p
37	57	89.1	13	23	ABJ11374	HLA binding antige
38	57	89.1	13	23	ABJ01954	158P1D7 related HL
39	57	89.1	13	23	ABG34860	Pan-DR-binding epi
40	57	89.1	13	23	AAU91547	Pan-DR-binding epi
41	57	89.1	13	23	ABB94471	Pan-DR-binding epi
42	57	89.1	13	23	AAU10848	Helper CD4 peptide
43	57	89.1	13	24	ABR01865	Pan-DR-binding epi
44	57	89.1	15	18	AAW22123	Padre (pan-DR bind
45	57	89.1	20	22	AAB46179	Tetanus toxoid epi

ALIGNMENTS

RESULT 1

AAR70250

ID AAR70250 standard; peptide; 13 AA.

XX

AC AAR70250;

XX

DT 25-MAR-2003 (updated)

DT 13-NOV-1995 (first entry)

XX

DE Pan DR-binding peptide activator of T cells.

XX

KW T cell; MHC; class II; major histocompatibility complex II;

KW immunogen; activation.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 1

FT /note= "D-form residue"

FT Misc-difference 13

FT /note= "D-form residue"

XX

PN WO9507707-A1.

XX

PD 23-MAR-1995.

XX

PF 14-SEP-1994; 94WO-US10368.

XX

PR 14-SEP-1993; 93US-0121101.

XX

PA (CYTE-) CYTEL CORP.

XX

PI Alexander JL, Gaeta FCA, Grey HM, Sette A, Sidney J;

XX

DR WPI; 1995-131178/17.

XX

PT Inhibiting or inducing an immune response using Pan DR-binding

PT peptide(s) - to prepare a compsn. used to treat allo:graft

PT rejection, allergic response and auto:immunity and as a vaccine

PT component

XX

PS Claim 3; Page 51; 59pp; English.

XX

CC Peptides capable of binding MHC class II molecules via antigen

CC binding sites (pan DR-binding peptides) were synthesised. Peptide

CC AAR70250 is capable of binding MHC class II molecules causing T cell

CC activation. This can be used as a vaccine component. The peptide can also

CC be used in combination with CTL peptides to enhance a CTL response.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 16; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00017;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13

Db

|||||
1 AKFVAAWTLKAAA 13

RESULT 2

AAW22121

ID AAW22121 standard; peptide; 13 AA.

XX

AC AAW22121;

XX

DT 13-MAR-1998 (first entry)

XX

DE Padre (pan-DR binding) peptide 6.

XX

KW PADRE; pan-DR binding; immune response; antigenic determinant;

KW treatment; tumour; infection.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 1

FT /note= "D-form residue"

FT Misc-difference 13

FT /note= "D-form residue"

XX

PN WO9726784-A1.

XX

PD 31-JUL-1997.

XX

PF 23-JAN-1997; 97WO-US01041.

XX

PR 24-JAN-1996; 96US-0010510.

XX

PA (CYTE-) CYTEL CORP.

XX

PI Alexander JL, Defrees S, Sette A;

XX

DR WPI; 1997-393272/36.

XX

PT Composition for eliciting immune response to non-protein determinant

PT - comprises the determinant and a pan-DR binding peptide, used for

PT prevention and treatment of tumours and infections

XX

PS Claim 23; Page 74; 87pp; English.

XX

CC This PADRE (pan-DR binding) peptide is used in a composition for
CC eliciting an immune response to a non-protein antigenic determinant.
CC The composition comprises of the PADRE peptide covalently linked to the
CC non-protein antigenic determinant. The composition is used to induce a
CC therapeutic or prophylactic response, particularly to selected
CC polysaccharide antigens associated with tumours or infectious agents. It
CC provides a high level, long-lasting IgG immune response. The composition
CC is also used to produce monoclonal antibodies which are potentially
CC useful as therapeutic and diagnostic agents. The composition can be used
CC to diagnose susceptibility of a patient to treatment with the non-protein
CC antigenic determinant or to predict subjects at risk from developing
CC chronic infections. PADRE peptides have broader specificity and higher

CC affinity than known DR-binding peptides. They are powerful inhibitors of
CC the proliferative response of human T cells restricted by at least 6
CC different DR molecules, and act as helper epitopes of in vivo induction
CC of cytotoxic T cells and antibody production.

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 18; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
|||
Db 1 AKFVAAWTLKAAA 13

RESULT 3

AAB36289

ID AAB36289 standard; Peptide; 13 AA.

XX

AC AAB36289;

XX

DT 26-FEB-2001 (first entry)

XX

DE Promiscuous T helper epitope SEQ ID NO: 65.

XX

KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW cancer; eosinophilia; vaccine; allergic rhinitis.

XX

OS Synthetic.

XX

PN WO200065058-A1.

XX

PD 02-NOV-2000.

XX

PF 19-APR-2000; 2000WO-DK00205.

XX

PR 23-APR-1999; 99DK-0000552.

PR 06-MAY-1999; 99US-0132811.

XX

PA (MEBI-) M & E BIOTECH AS.

XX

PI Klysner S;

XX

DR WPI; 2000-672791/65.

XX

PT Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -

XX

PS Disclosure; Page 168; 172pp; English.

XX

CC The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and

CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections.

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13

|||||||

Db 1 AKFVAAWTLKAAA 13

RESULT 4

AAAY52558

ID AAY52558 standard; peptide; 13 AA.

XX

AC AAY52558;

XX

DT 28-FEB-2000 (first entry)

XX

DE Universal helper T epitope, pan DR epitope (PADRE).

XX

KW Chimeric; Ii protein; pan DR epitope; expression vector;
KW promoter; major histocompatibility complex; MHC; targetting; peptide;
KW epitope; antigen; presentation; class I; cytosolic pathway;
KW endoplasmic reticulum; class II; extracellular antigen;
KW endocytic pathway; helper T lymphocyte; HTL; universal epitope;
KW cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;
KW vaccine; immunity; infection; pathogen; virus; HIV; HBV; HCV;
KW hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;
KW autoimmune disease; activation; antiviral; antimalarial;
KW immunoprotective.

XX

OS Synthetic.

XX

PN WO9958658-A2.

XX

PD 18-NOV-1999.

XX

PF 13-MAY-1999; 99WO-US10646.

XX

PR 13-MAY-1998; 98US-0078904.

PR 15-MAY-1998; 98US-0085751.

XX

PA (EPIM-) EPIMMUNE INC.

XX

PI Fikes JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;
PI Chesnut RW;

XX

DR WPI; 2000-039103/03.

DR N-PSDB; AAZ38679.

XX

PT Expression vectors encoding major histocompatibility targeting

PT sequence, used as, e.g. tumor vaccines -

XX

PS Claim 9; Page 80; 130pp; English.

XX

CC This sequence represents a universal helper T epitope, pan DR epitope
CC (PADRE), DNA encoding which is used to construct fusion genes used
CC in exemplifications of the present invention. The invention
CC relates to a novel expression vector comprising a promoter operably
CC linked to a fusion gene encoding a major histocompatibility complex
CC (MHC) targetting sequence, and two or more heterologous peptide epitopes.
CC The MHC targetting sequence may be a class I targetting sequence, which
CC directs an MHC class I epitope to a cytosolic pathway or to the
CC endoplasmic reticulum, or an MHC class II targetting sequence, which
CC directs extracellular antigens to enter the endocytic pathway to be
CC processed into antigen peptides for presentation on MHC class II
CC molecules. The heterologous epitopes may comprise either helper T
CC lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL) epitope and
CC a universal HTL epitope such as a pan DR epitope (PADRE). The vectors are
CC useful for stimulating an immune response in vivo, as well as for use in
CC assaying the human immunogenicity of a human T cell peptide epitope in
CC vivo in a non-human mammal. They provide a nucleic acid vaccine for
CC enhancing immunity against infectious pathogens, such as viruses (e.g.,
CC HIV, hepatitis B (HBV and hepatitis C (HCV)) bacteria, protozoa (e.g.,
CC Plasmodium falciparum, the cause of malaria) and also tumour cells and
CC autoimmune diseases. Universal MHC class II epitopes are advantageously
CC combined with other MHC class I and class II epitopes to increase the
CC number of cells that are activated in response to a given antigen and
CC provide a broader population coverage of MHC-reactive alleles.

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 21; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00017;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13

|||||||

Db 1 AKFVAAWTLKAAA 13

RESULT 5

AAB73644

ID AAB73644 standard; peptide; 13 AA.

XX

AC AAB73644;

XX

DT 11-SEP-2001 (first entry)

XX

DE Pan-DR binding peptide, PADRE.

XX

KW Cellular vaccine; antigen-presenting cell; APC;
KW cell-surface molecule density; major histocompatibility complex; MHC;
KW antigen-specific T-cell expansion; tumour; cancer; viral infection;
KW parasitic infection; Pan-DR binding peptide; PADRE.

XX

OS Unidentified.

XX

PN WO200136978-A1.
 XX
 PD 25-MAY-2001.
 XX
 PF 15-NOV-2000; 2000WO-US42213.
 XX
 PR 15-NOV-1999; 99US-0165428.
 XX
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Schultze JL, Vonderheide RH, Nadler LM, Maecker B;
 PI Von Bergwelt-Baildon M;
 XX
 DR WPI; 2001-343909/36.
 XX
 PT Determining surface density of molecules on antigen-presenting cells,
 PT useful for quality control of cellular vaccines for treatment of tumors
 PT and infections comprises measuring density ratios -
 XX
 PS Disclosure; Page 21; 47pp; English.
 XX
 CC The invention relates to a method of determining the surface density of
 CC a cell-surface molecule on a primary or artificial antigen-presenting
 CC cell (APC). The method comprises determining the cell surface area
 CC of the APCs, determining the absolute amount of cell surface molecule on
 CC the surface of the APCs, and calculating the ratio of the amount of
 CC the cell surface molecule to the APC surface area as a measure of cell
 CC surface molecule density. The invention also encompasses a kit for
 CC determining the cell-surface density of a complex of peptide and MHC
 CC (major histocompatibility complex) protein/peptide complex on an APC;
 CC identifying an agent that increases persistence of the MHC/peptide
 CC complex on an APC surface by culturing APCs in presence of test compound
 CC and measuring the time of persistence relative to an untreated control;
 CC and determining if the APC carries a therapeutically adequate amount of
 CC peptide by determining whether the cell-surface density of the
 CC MHC/peptide complex is 100 molecules/square micrometre or more. The
 CC method is used to determine if the amount of MHC/antigenic peptide
 CC complex present on the surface of an APC is sufficient for therapeutic
 CC use of the cells as cellular vaccines or for ex vivo expansion of
 CC antigen-specific T cells for subsequent return to the patient,
 CC particularly for the treatment of tumours or viral or parasitic
 CC infections. The method can also be used to identify agents that increase
 CC production of MHC/peptide complexes on cells, (which increases the
 CC therapeutic potential of the cells) by treatment before and/or during
 CC APC administration. The method allows identification of therapeutically
 CC useful antigen-pulsed APC optimisation of conditions for their
 CC production, and control of APC quality. Sequences AAB73643-AAB73648
 CC represent peptides used in MHC binding studies and for the generation
 CC of peptide-specific cytotoxic T-lymphocytes.
 XX
 SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 22; Length 13;
 Best Local Similarity 100.0%; Pred. No. 0.00017;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13

Db

|||||
1 AKFVAAWTLKAAA 13

RESULT 6

AAB99710

ID AAB99710 standard; peptide; 13 AA.

XX

AC AAB99710;

XX

DT 06-SEP-2001 (first entry)

XX

DE Pan-DR-binding peptide (PADRE) SEQ ID NO:31.

XX

KW Human leukocyte antigen A2 binding peptide; HLA class I A2; CTL;
KW cytotoxic T-cell lymphocyte; tumour associated antigen; CEA; HER2/neu;
KW MAGE2; MAGE3; p53; vaccine; cancer; cytostatic; immunomodulator;
KW immunotherapy; immune response.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200141741-A1.

XX

PD 14-JUN-2001.

XX

PF 13-DEC-2000; 2000WO-US34318.

XX

PR 13-DEC-1999; 99US-0170448.

PR 05-APR-2000; 2000US-0543608.

PR 30-MAY-2000; 2000US-0583200.

XX

PA (EPIM-) EPIMMUNE INC.

XX

PI Fikes J, Sette A, Sidney J, Southwood S, Celis E, Keogh E;
PI Chesnut R;

XX

DR WPI; 2001-381489/40.

XX

PT Compositions for use in a vaccine for treating, e.g., breast, lung and
PT colon cancer comprises at least one peptide that comprises an isolated
PT epitope of a tumor-associated antigen -

XX

PS Example 7; Page 48; 86pp; English.

XX

CC The present invention describes a composition (I) comprising at least
CC one peptide that comprises an isolated, prepared epitope consisting of
CC a sequence selected from 25 short amino acid sequences given in AAB99680
CC to AAB99704. Also described are: (1) a composition (II) comprising one
CC or more peptides, and further comprising at least two epitopes selected
CC from the 25 short amino acid sequences (as above), where each of the one
CC or more peptides comprise less than 50 contiguous amino acids that have
CC 100% identity with a native peptide sequence; and (2) a vaccine
CC composition (III) comprising an epitope selected from the 25 short amino
CC acid sequences (as above) and a pharmaceutical excipient. (I) has
CC cytostatic and immunomodulatory activities and can be used in vaccine
CC production and immunotherapy. The peptide epitope compositions (I)-(II)

CC are useful for monitoring an immune response to a tumour associated
CC antigen or when one or more peptides are combined to create a vaccine
CC (III) that stimulates the cellular arm of the immune system. In
CC particular, the vaccine mediates immune responses against tumours in
CC individuals who bear an allele of the human leukocyte antigen (HLA)-A2
CC supertype and improve the standard of care for patients being treated
CC for breast, colon, or lung cancer. The present sequence represents a
CC pan-DR-binding peptide (PADRE) sequence, which is used in an example
CC from the present invention.

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
|||
Db 1 AKFVAAWTLKAAA 13

RESULT 7

AAB20154

ID AAB20154 standard; Peptide; 13 AA.

XX

AC AAB20154;

XX

DT 30-APR-2001 (first entry)

XX

DE PADRE peptide.

XX

KW Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;

KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;

KW cardiant; PADRE; pan DE epitope.

XX

OS Synthetic.

XX

PN WO200105820-A2.

XX

PD 25-JAN-2001.

XX

PF 20-JUL-2000; 2000WO-DK00413.

XX

PR 20-JUL-1999; 99DK-0001014.

PR 26-JUL-1999; 99US-0145275.

XX

PA (MEBI-) M & E BIOTECH AS.

XX

PI Halkier T, Mouritsen S, Klysner S;

XX

DR WPI; 2001-112680/12.

XX

PT Increasing the muscle mass of animals used in meat production by down

PT regulating growth differentiation factor 8 (GDF-8) activity in the

PT animal through induction of anti-GDF-8 antibody production -

XX

PS Disclosure; Page 15; 110pp; English.

XX

CC The present sequence is that of a PADRE (pan DE epitope) peptide
CC which acts as a T-cell epitope and is capable of binding a large
CC proportion of major histocompatibility complex class II molecules.
CC It is an object of the invention to produce a recombinant
CC therapeutic vaccine that is capable of effecting down-regulation of
CC growth differentiation factor 8 (GDF-8) in order to increase the
CC muscle growth rate of farm animals. The vaccines (see AAB20145-53)
CC are capable of breaking autotolerance against autologous GDF-8.
CC They comprise the C-terminal portion of human GDF-8 in which a
CC portion of the native sequence is replaced by a T-cell epitope such
CC as PADRE, with minimal disturbance of the authentic 3-dimensional
CC structure of the protein. Down-regulation of GDF-8 activity can
CC increase muscle mass by up to at least 45% in cattle, pigs and
CC poultry used for meat production, reducing the need for antibiotic
CC feed-additives. Anti-GDF8 vaccines can be used to treat human
CC diseases such as cancer cachexia where muscle atrophy is pronounced
CC and for patients suffering from acute and chronic heart failure.

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13

|||||||

Db 1 AKFVAAWTLKAAA 13

RESULT 8

AAE26369

ID AAE26369 standard; peptide; 13 AA.

XX

AC AAE26369;

XX

DT 13-DEC-2002 (first entry)

XX

DE PADRE peptide.

XX

KW Human; immune response; T-helper cell epitope; chitosan; CTL response;

KW vaccine; prostate cancer; breast cancer; cytostatic; immunostimulant.

XX

OS Unidentified.

XX

PN WO200234287-A2.

XX

PD 02-MAY-2002.

XX

PF 26-OCT-2001; 2001WO-DK00705.

XX

PR 27-OCT-2000; 2000DK-0001606.

PR 03-NOV-2000; 2000US-245166P.

PR 18-JUN-2001; 2001DK-0000936.

XX

PA (PHAR-) PHARMEXA AS.

XX

PI Beier AM, Gautam A, Mouritsen S;
XX
DR WPI; 2002-463339/49.
XX
PT Inducing or enhancing an immune response against an antigen,
PT particularly cytotoxic T-lymphocyte responses, for treating or
PT ameliorating prostate or breast cancer, comprises administering the
PT antigen formulated with chitosan -
XX

PS Disclosure; Page 32; 97pp; English.
XX

CC The invention relates to a method for inducing or enhancing an immune
CC response against a polypeptide antigen in an animal, including human.
CC The method comprises administering the polypeptide antigen or at least
CC one variant which includes at least one first T-helper cell epitope that
CC is foreign to the animal (foreign TH epitope) and is formulated with
CC chitosan. The polypeptide antigen is weakly immunogenic or non-
CC immunogenic. The invention is used as vaccine. The chitosan and
CC polypeptide antigen or its variant are useful in the preparation of an
CC immunogenic composition for inducing or enhancing an immune response,
CC particularly CTL response, against the polypeptide or protein antigen.
CC The method for inducing or enhancing an immune response is useful in
CC treating or ameliorating cancer, e.g. prostate or breast cancer. The
CC present sequence is PADRE peptide used to illustrate the method
CC of the invention.
XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13
|||
Db 1 AKFVAAWTLKAAA 13

RESULT 9

ABG31776

ID ABG31776 standard; Peptide; 13 AA.

XX

AC ABG31776;

XX

DT 03-DEC-2002 (first entry)

XX

DE Pan DR epitope (PADRE) peptide.

XX

KW Immunogen; B-cell epitope; cytotoxic T lymphocyte; CTL; TH epitope;

KW T helper cell epitope; virtual lymph node device; pan DR epitope; PADRE.

XX

OS Synthetic.

XX

PN WO200266056-A2.

XX

PD 29-AUG-2002.

XX

PF 19-FEB-2002; 2002WO-DK00112.

XX
 PR 19-FEB-2001; 2001WO-DK00113.
 PR 20-FEB-2001; 2001US-0785215.
 PR 20-AUG-2001; 2001DK-0001231.
 PR 22-OCT-2001; 2001US-337543P.
 XX
 PA (PHAR-) PHARMEXA AS.
 XX
 PI Nielsen KG, Koefoed P;
 XX
 DR WPI; 2002-706932/76.
 XX
 PT Novel immunogen useful for immunising an animal, has an activated
 PT polyhydroxypolymer backbone to which is attached an antigenic
 PT determinant including a B cell epitope and another determinant
 PT including a T-helper epitope -
 XX
 PS Disclosure; Page 17; 52pp; English.
 XX
 CC The invention relates to an immunogen comprising at least one first
 CC antigenic determinant that includes at least one B-cell epitope and/or at
 CC least one cytotoxic T lymphocyte (CTL) epitope, and at least one second
 CC antigenic determinant that includes a T helper cell epitope (TH epitope),
 CC where each of the first and second antigenic determinants are coupled to
 CC an activated polyhydroxypolymer carrier. The invention also relates to an
 CC immunogenic composition for raising an immune response against an antigen
 CC in a mammal, including a human. The immunogen or immunogenic composition
 CC contained in a virtual lymph node (VLN) device is useful for immunising
 CC an animal, including a human, against an antigen of choice, where the
 CC antigen shares at least one first antigenic determinant with the
 CC immunogen. This sequence represents a pan DR epitope (PQDRE) peptide
 CC which is a T helper cell epitope used in synthesis of an immunogen of the
 CC invention.
 XX
 SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 23; Length 13;
 Best Local Similarity 100.0%; Pred. No. 0.00017;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13
 |||||
 Db 1 AKFVAAWTLKAAA 13

RESULT 10

AAU80293

ID AAU80293 standard; Peptide; 13 AA.

XX

AC AAU80293;

XX

DT 30-JUL-2002 (first entry)

XX

DE Pan DR epitope peptides (PADRE) epitope #2.

XX

KW IgE; allergy; antiallergic; immunosuppressive; antianaphylactic;
 KW antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IgE;

KW vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;
KW heavy chain C domain; PADRE; Pan DR epitope peptide.
XX
OS Synthetic.
XX
PN WO200220038-A2.
XX
PD 14-MAR-2002.
XX
PF 06-SEP-2001; 2001WO-DK00579.
XX
PR 06-SEP-2000; 2000DK-0001326.
PR 15-SEP-2000; 2000US-232831P.
XX
PA (PHAR-) PHARMEXA AS.
XX
PI Klysner S, Von Hoegen P, Voldborg B, Gautam A;
XX
DR WPI; 2002-383033/41.
DR N-PSDB; ABK51139.
XX
PT Inducing immune response against autologous immunoglobulin E in an
PT animal, by effecting simultaneous presentation of cytotoxic T
PT lymphocyte epitope an/or B-cell epitope derived from the immunoglobulin
PT -
XX
PS Disclosure; Page 31; 151pp; English.
XX
CC This invention relates to a novel method for inducing an immune response
CC against autologous immunoglobulin E (IgE) in an animal. The method
CC comprises effecting simultaneous presentation of cytotoxic T lymphocyte
CC (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell
CC epitope (TH epitope) which is foreign to the animal, by antigen
CC presenting cells (APCs) of the animal's immune system. The epitopes
CC of the invention may be used as a vaccine against allergic diseases. The
CC method of the invention is useful for inducing an immune response
CC against autologous IgE in an animal, which is useful for downregulating
CC autologous IgE in the animal. This method is useful in the prevention
CC and treatment of allergic diseases such as anaphylaxis, allergic
CC rhinitis, asthma and atopic dermatitis. The present sequence represents
CC human Pan DR epitope (PADRE) #2 used in the method of the invention.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
|||||||
Db 1 AKFVAAWTLKAAA 13

RESULT 11
ABP72696
ID ABP72696 standard; Peptide; 13 AA.
XX

AC ABP72696;
 XX
 DT 11-JUN-2003 (first entry)
 XX
 DE T cell epitope PADRE.
 XX
 KW PADRE; epitope; amyloid precursor protein; APP; beta amyloid;
 KW vaccine; genetic immunisation; nootropic; neuroprotective;
 KW Alzheimer's disease.
 XX
 OS Synthetic.
 XX
 PN WO2003015812-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 20-AUG-2002; 2002WO-DK00547.
 XX
 PR 20-AUG-2001; 2001DK-0001231.
 PR 22-OCT-2001; 2001US-337543P.
 PR 16-APR-2002; 2002DK-0000558.
 PR 16-APR-2002; 2002US-373027P.
 XX
 PA (PHAR-) PHARMEXA AS.
 XX
 PI Rasmussen PB, Jensen MR, Nielsen KG, Koefoed P, Degan FD;
 XX
 DR WPI; 2003-312718/30.
 XX
 PT Novel analog of amyloid precursor protein or beta amyloid for treating
 PT Alzheimer's disease, has amyloid precursor protein/beta amyloid
 PT incorporating B-cell epitope of amyloid protein and foreign T-helper
 PT epitope -
 XX
 PS Disclosure; Page 38; 122pp; English.
 XX
 CC The present sequence is that of the pan DR epitope (PADRE), an
 CC artificial T cell epitope which is capable of binding a large
 CC proportion of MHC Class II molecules. The invention provides
 CC methods for compositions for combatting diseases characterised by
 CC deposition of amyloid, such as Alzheimer's disease. Immunisation
 CC is preferably effected by administration of analogues of autologous
 CC amyloid precursor protein (APP) or beta amyloid (Abeta), the
 CC analogues being capable of inducing antibody production against
 CC the autologous amyloidogenic polypeptides. Especially preferred
 CC as an immunogen is autologous Abeta which has been modified by
 CC introduction of one or a few foreign, immunodominant and
 CC promiscuous T-cell epitopes, such as PADRE. Genetic immunisation
 CC against APP or Abeta and vaccination using live vaccines are also
 CC provided.
 XX
 SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 24; Length 13;
 Best Local Similarity 100.0%; Pred. No. 0.00017;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
|||||||
Db 1 AKFVAAWTLKAAA 13

RESULT 12

AAU80290

ID AAU80290 standard; Protein; 55 AA.

XX

AC AAU80290;

XX

DT 30-JUL-2002 (first entry)

XX

DE Pan DR epitope peptides (PADRE) epitope.

XX

KW IgE; allergy; antiallergic; immunosuppressive; antianaphylactic;
KW antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IgE;
KW vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;
KW heavy chain C domain; PADRE; Pan DR epitope peptide.

XX

OS Synthetic.

XX

PN WO200220038-A2.

XX

PD 14-MAR-2002.

XX

PF 06-SEP-2001; 2001WO-DK00579.

XX

PR 06-SEP-2000; 2000DK-0001326.

PR 15-SEP-2000; 2000US-232831P.

XX

PA (PHAR-) PHARMEXA AS.

XX

PI Klysner S, Von Hoegen P, Voldborg B, Gautam A;

XX

DR WPI; 2002-383033/41.

DR N-PSDB; ABK51136.

XX

PT Inducing immune response against autologous immunoglobulin E in an
PT animal, by effecting simultaneous presentation of cytotoxic T
PT lymphocyte epitope an/or B-cell epitope derived from the immunoglobulin

PT -

XX

PS Disclosure; Page 126; 151pp; English.

XX

CC This invention relates to a novel method for inducing an immune response
CC against autologous immunoglobulin E (IgE) in an animal. The method
CC comprises effecting simultaneous presentation of cytotoxic T lymphocyte
CC (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell
CC epitope (TH epitope) which is foreign to the animal, by antigen
CC presenting cells (APCs) of the animal's immune system. The epitopes
CC of the invention may be used as a vaccine against allergic diseases. The
CC method of the invention is useful for inducing an immune response
CC against autologous IgE in an animal, which is useful for downregulating
CC autologous IgE in the animal. This method is useful in the prevention
CC and treatment of allergic diseases such as anaphylaxis, allergic
CC rhinitis, asthma and atopic dermatitis. The present sequence represents

CC a human Pan DR epitope (PADRE) used in the method of the invention.

XX

SQ Sequence 55 AA;

Query Match 100.0%; Score 64; DB 23; Length 55;

Best Local Similarity 100.0%; Pred. No. 0.00082;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13

|||||||

Db 13 AKFVAAWTLKAAA 25

RESULT 13

AAY52548

ID AAY52548 standard; Protein; 80 AA.

XX

AC AAY52548;

XX

DT 28-FEB-2000 (first entry)

XX

DE Murine Ig-kappa signal/multiple MHC class II epitope fusion protein.

XX

KW Chimeric; Ig kappa; pan DR epitope; expression vector;

KW promoter; major histocompatibility complex; MHC; targetting; peptide;

KW epitope; antigen; presentation; class I; cytosolic pathway;

KW endoplasmic reticulum; class II; extracellular antigen;

KW endocytic pathway; helper T lymphocyte; HTL; universal epitope;

KW cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;

KW vaccine; immunity; infection; pathogen; virus; HIV; HBV; HCV;

KW hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;

KW autoimmune disease; activation; antiviral; antimalarial;

KW immunoprotective; immunoglobulin.

XX

OS Synthetic.

OS Mus sp.

XX

PN WO9958658-A2.

XX

PD 18-NOV-1999.

XX

PF 13-MAY-1999; 99WO-US10646.

XX

PR 13-MAY-1998; 98US-0078904.

PR 15-MAY-1998; 98US-0085751.

XX

PA (EPIM-) EPIMMUNE INC.

XX

PI Fikes JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;

PI Chesnut RW;

XX

DR WPI; 2000-039103/03.

DR N-PSDB; AAZ38648.

XX

PT Expression vectors encoding major histocompatibility targeting

PT sequence, used as, e.g. tumor vaccines -

XX

PS Example 1; Fig 11; 130pp; English.

XX

CC This sequence represents the fusion protein encoded
CC by a murine Ig kappa signal sequence/multiple T helper
CC epitope fusion gene, SigTh, used in an exemplification of the
CC present invention. The invention relates to a novel expression vector
CC comprising a promoter operably linked to a fusion gene encoding a major
CC histocompatibility complex (MHC) targetting sequence, and two or more
CC heterologous peptide epitopes. The MHC targetting sequence may be a class
CC I targetting sequence, which directs an MHC class I epitope to a
CC cytosolic pathway or to the endoplasmic reticulum, or an MHC class II
CC targetting sequence, which directs extracellular antigens to enter
CC the endocytic pathway to be processed into antigen peptides for
CC presentation on MHC class II molecules. The heterologous epitopes
CC may comprise either helper T lymphocyte (HTL) epitopes, or a
CC cytotoxic T lymphocyte (CTL) epitope and a universal HTL epitope
CC such as a pan DR epitope (PADRE). The vectors are useful for
CC stimulating an immune response in vivo, as well as for use in
CC assaying the human immunogenicity of a human T cell peptide epitope in
CC vivo in a non-human mammal. They provide a nucleic acid vaccine for
CC enhancing immunity against infectious pathogens, such as viruses (e.g.,
CC HIV, hepatitis B (HBV and hepatitis C (HCV)) bacteria, protozoa (e.g.,
CC Plasmodium falciparum, the cause of malaria) and also tumour cells and
CC autoimmune diseases. Universal MHC class II epitopes are advantageously
CC combined with other MHC class I and class II epitopes to increase the
CC number of cells that are activated in response to a given antigen and
CC provide a broader population coverage of MHC-reactive alleles.

XX

SQ Sequence 80 AA;

Query Match 100.0%; Score 64; DB 21; Length 80;

Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13

|||||||

Db 68 AKFVAAWTLKAAA 80

RESULT 14

AA52541

ID AAY52541 standard; Protein; 118 AA.

XX

AC AAY52541;

XX

DT 28-FEB-2000 (first entry)

XX

DE KappaLAMP-Th fusion protein.

XX

KW Chimeric; Ig-kappa; LAMP-1; pan DR epitope; expression vector;
KW promoter; major histocompatibility complex; MHC; targetting; peptide;
KW epitope; antigen; presentation; class I; cytosolic pathway;
KW endoplasmic reticulum; class II; extracellular antigen;
KW endocytic pathway; helper T lymphocyte; HTL; universal epitope;
KW cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;
KW vaccine; immunity; infection; pathogen; virus; HIV; HBV; HCV;
KW hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;

KW autoimmune disease; activation; antiviral; antimalarial;
 KW immunoprotective; immunoglobulin.
 XX
 OS Synthetic.
 OS Mus sp.
 XX
 PN WO9958658-A2.
 XX
 PD 18-NOV-1999.
 XX
 PF 13-MAY-1999; 99WO-US10646.
 XX
 PR 13-MAY-1998; 98US-0078904.
 PR 15-MAY-1998; 98US-0085751.
 XX
 PA (EPIM-) EPIMMUNE INC.
 XX
 PI Fikes JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;
 PI Chesnut RW;
 XX
 DR WPI; 2000-039103/03.
 DR N-PSDB; AAZ38619.
 XX
 PT Expression vectors encoding major histocompatibility targeting
 PT sequence, used as, e.g. tumor vaccines -
 XX
 PS Example 1; Fig 4; 130pp; English.
 XX
 CC This sequence represents the fusion protein encoded
 CC by a murine Ig-kappa signal sequence/multiple T helper
 CC epitope/lysosomal membrane glycoprotein-1 (LAMP-1) fusion gene,
 CC KappaLAMP-Th, used in an exemplification of the present invention. The
 CC invention relates to a novel expression vector comprising a promoter
 CC operably linked to a fusion gene encoding a major histocompatibility
 CC complex (MHC) targetting sequence, and two or more heterologous peptide
 CC epitopes. The MHC targetting sequence may be a class I targetting
 CC sequence, which directs an MHC class I epitope to a cytosolic pathway or
 CC to the endoplasmic reticulum, or an MHC class II targetting sequence,
 CC which directs extracellular antigens to enter the endocytic pathway to be
 CC processed into antigen peptides for presentation on MHC class II
 CC molecules. The heterologous epitopes may comprise either helper T
 CC lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL) epitope and
 CC a universal HTL epitope such as a pan DR epitope (PADRE). The vectors are
 CC useful for stimulating an immune response in vivo, as well as for use in
 CC assaying the human immunogenicity of a human T cell peptide epitope in
 CC vivo in a non-human mammal. They provide a nucleic acid vaccine for
 CC enhancing immunity against infectious pathogens, such as viruses (e.g.,
 CC HIV, hepatitis B (HBV and hepatitis C (HCV)) bacteria, protozoa (e.g.,
 CC Plasmodium falciparum, the cause of malaria) and also tumour cells and
 CC autoimmune diseases. Universal MHC class II epitopes are advantageously
 CC combined with other MHC class I and class II epitopes to increase the
 CC number of cells that are activated in response to a given antigen and
 CC provide a broader population coverage of MHC-reactive alleles.
 XX
 SQ Sequence 118 AA;

Query Match

100.0%; Score 64; DB 21; Length 118;

Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13
|||
Db 68 AKFVAAWTLKAAA 80

RESULT 15

AAAY52542

ID AAY52542 standard; Protein; 119 AA.

XX

AC AAY52542;

XX

DT 28-FEB-2000 (first entry)

XX

DE Murine H2-M/multiple MHC class II epitope fusion protein.

XX

KW Chimeric; H2-M; pan DR epitope; expression vector;
KW promoter; major histocompatibility complex; MHC; targetting; peptide;
KW epitope; antigen; presentation; class I; cytosolic pathway;
KW endoplasmic reticulum; class II; extracellular antigen;
KW endocytic pathway; helper T lymphocyte; HTL; universal epitope;
KW cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;
KW vaccine; immunity; infection; pathogen; virus; HIV; HBV; HCV;
KW hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;
KW autoimmune disease; activation; antiviral; antimalarial;
KW immunoprotective.

XX

OS Synthetic.

OS Mus sp.

XX

PN WO9958658-A2.

XX

PD 18-NOV-1999.

XX

PF 13-MAY-1999; 99WO-US10646.

XX

PR 13-MAY-1998; 98US-0078904.

PR 15-MAY-1998; 98US-0085751.

XX

PA (EPIM-) EPIMMUNE INC.

XX

PI Fikes JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;
PI Chesnut RW;

XX

DR WPI; 2000-039103/03.

DR N-PSDB; AAZ38620.

XX

PT Expression vectors encoding major histocompatibility targeting
PT sequence, used as, e.g. tumor vaccines -

XX

PS Example 1; Fig 5; 130pp; English.

XX

CC This sequence represents the fusion protein encoded by
CC a murine H2-M/multiple MHC class II epitope fusion gene,
CC H2M-Th, used in an exemplification of the present invention.

CC The invention relates to a novel expression vector comprising a promoter
CC operably linked to a fusion gene encoding a major histocompatibility
CC complex (MHC) targetting sequence, and two or more heterologous peptide
CC epitopes. The MHC targetting sequence may be a class I targetting
CC sequence, which directs an MHC class I epitope to a cytosolic pathway or
CC to the endoplasmic reticulum, or an MHC class II targetting sequence,
CC which directs extracellular antigens to enter the endocytic pathway to be
CC processed into antigen peptides for presentation on MHC class II
CC molecules. The heterologous epitopes may comprise either helper T
CC lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL) epitope and
CC a universal HTL epitope such as a pan DR epitope (PADRE). The vectors are
CC useful for stimulating an immune response in vivo, as well as for use in
CC assaying the human immunogenicity of a human T cell peptide epitope in
CC vivo in a non-human mammal. They provide a nucleic acid vaccine for
CC enhancing immunity against infectious pathogens, such as viruses (e.g.,
CC HIV, hepatitis B (HBV and hepatitis C (HCV)) bacteria, protozoa (e.g.,
CC Plasmodium falciparum, the cause of malaria) and also tumour cells and
CC autoimmune diseases. Universal MHC class II epitopes are advantageously
CC combined with other MHC class I and class II epitopes to increase the
CC number of cells that are activated in response to a given antigen and
CC provide a broader population coverage of MHC-reactive alleles.

XX

SQ Sequence 119 AA;

Query Match 100.0%; Score 64; DB 21; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13
|||
Db 63 AKFVAAWTLKAAA 75

Search completed: December 13, 2003, 13:34:24
Job time : 3.15873 secs